

12-1-2011

A Comparison of Current Anuran Monitoring Methods with Emphasis on the Accuracy of Automatic Vocalization Detection Software

Jacob Douglas Eldridge

Western Kentucky University, jacob.eldridge518@topper.wku.edu

Follow this and additional works at: <http://digitalcommons.wku.edu/theses>



Part of the [Biodiversity Commons](#), and the [Population Biology Commons](#)

Recommended Citation

Eldridge, Jacob Douglas, "A Comparison of Current Anuran Monitoring Methods with Emphasis on the Accuracy of Automatic Vocalization Detection Software" (2011). *Masters Theses & Specialist Projects*. Paper 1122.
<http://digitalcommons.wku.edu/theses/1122>

This Thesis is brought to you for free and open access by TopSCHOLAR®. It has been accepted for inclusion in Masters Theses & Specialist Projects by an authorized administrator of TopSCHOLAR®. For more information, please contact connie.foster@wku.edu.

A COMPARISON OF CURRENT ANURAN MONITORING METHODS WITH
EMPHASIS ON THE ACCURACY OF AUTOMATIC VOCALIZATION
DETECTION SOFTWARE

A Thesis
Presented to
The Faculty of the Department of Biology
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

By
Jacob Douglas Eldridge


December 2011

A COMPARISON OF CURRENT ANURAN MONITORING METHODS WITH EMPHASIS
ON THE ACCURACY OF AUTOMATIC VOCALIZATION DETECTION SOFTWARE

Date Recommended November 19, 2011



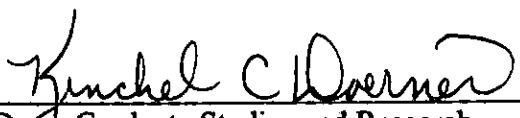
Dr. José Pedro do Amaral, Director of Thesis



Dr. Richard G. Bowker



Dr. Albert Meier



Dear, Graduate Studies and Research 6-Jan-2012
Date

ACKNOWLEDGMENTS

I would like to thank the faculty, staff, and students at Western Kentucky University for a great graduate education experience. Particularly, I thank my advisor, Dr. José Pedro do Amaral, and committee members, Drs. Richard G. Bowker and Albert Meier, for their ideas, support, and guidance in my thesis research.

I would also like to acknowledge the support of the organizations that made this study possible. I am grateful for the funding I received from Ministerio de Ciencia e Innovación, Spain (CGL2008-04814-C02-01), Western Kentucky University Graduate Studies, Western Kentucky University Biology Department, USDA Natural Resource Conservation Service, NRCS, and the Western Kentucky University Upper Green River Biological Preserve that allowed this project to become a reality. Additionally, the management of the Upper Green River Biological Preserve was gracious in providing a beautiful and unique location for this study.

There are numerous people that deserve thanks for their role in this research. I certainly appreciate the numerous hours spent by the Honors Ecology students listening to audio recordings. These students include Shouta Brown, Taylor R. Bryant, Rachel Calhoun, Clarice Esch, Ciera Gary, Jessica King, Kara L. McCarthy, Ashley M. Mefford, Viktoria E. Nelin, Mary Newton, Cheryl C. Onwu, Delaney Rockrohr, Andrea Sejdic, Brenna E. Tinsley, Hope Tucker, Lindsay Williams, and Maggie Wilder. I would also like to thank Jeanne Shearer, Derek Rupert, Ashley Mefford, Maggie Wilder, Viktoria Nelin, Angela Eldridge, and Nick Basham for their help in various aspects of field surveys. I am also grateful to Nick Basham and Emily Hollis for servicing the Song Meters. The ideas and editing from fellow graduate students Maggie Hook, Derek

Rupert, and Maggie Wisniewska were also valuable. I am indebted to Drs. Ouida Meier and Mike Collyer for additional help with statistical methods.

Finally, I am most thankful for the support of my wife, Angela Eldridge, who has physically, mentally, and emotionally helped me to complete this thesis. My family has also provided the direction and push to do what I love. I am very grateful to all who have encouraged me to pursue my passion to preserve earth's biodiversity.

TABLE OF CONTENTS

List of Figures.....	vi
List of Tables.....	ix
Abstract.....	xi

Chapter 1

Abstract.....	1
Introduction.....	2
Methods.....	7
Results.....	27
Discussion.....	49
Appendix 1.....	66
Bibliography.....	72

Chapter 2

Abstract.....	77
Introduction.....	79
Methods.....	82
Results.....	91
Discussion.....	108
Bibliography.....	120

LIST OF FIGURES

Figure 1.1 Aerial photograph of the two digital recorder sites at Weldon Peete Park in Bowling Green, KY used for bioacoustic monitoring in the study.....	10
Figure 1.2 Aerial photograph of the two digital recorder sites at the Upper Green River Biological Preserve in Hart County, KY used for bioacoustic monitoring in the study.....	11
Figure 1.3 A schematic of the site used for physical and auditory surveys including the relative locations of the two fields and the survey methods (drift fences, poly-vinyl chloride (PVC) pipe array, transects, and recorders).....	17
Figure 1.4 Histogram comparison of the total monthly species richness as detected by all methods used in this study.....	33
Figure 1.5 Line plot comparison of total daily detection of <i>Bufo fowleri</i> in its breeding season using physical and auditory methods.....	35
Figure 1.6 Line plot comparison of total weekly detection of <i>Bufo fowleri</i> in its breeding season using physical and auditory methods.....	36
Figure 1.7 Line plot comparison of <i>Hyla chrysoscelis</i> detections by new individuals captured, recaptured individuals, and auditory detection presences.....	37
Figure 1.8 Histogram comparison of the proportion of calling presence during each period of the day of <i>Bufo americanus</i> at the Upper Green River Biological Preserve and Weldon Peete Park.....	38
Figure 1.9 Histogram comparison of the proportion of calling presence during each period of the day of <i>Hyla chrysoscelis</i> at the Upper Green River Biological Preserve and Weldon Peete Park.....	39
Figure 1.10 Histogram comparison of the proportion of calling presence during each period of the day of <i>Pseudacris crucifer</i> at the Upper Green River Biological Preserve and Weldon Peete Park.....	40
Figure 1.11 Dates of detected calls of <i>Bufo americanus</i> at the Upper Green River Biological Preserve and Weldon Peete Park.....	41
Figure 1.12 Dates of detected calls of <i>Hyla chrysoscelis</i> at the Upper Green River Biological Preserve and Weldon Peete Park.....	42
Figure 1.13 Dates of detected calls of <i>Pseudacris crucifer</i> at the Upper Green River Biological Preserve and Weldon Peete Park.....	43

Figure 1.14 Boxplot comparison of the call length of <i>Hyla chrysoscelis</i> at both study sites, Weldon Peete Park and the Upper Green River Biological Preserve.....	47
Figure 1.15 Boxplot comparison of the call length of <i>Pseudacris crucifer</i> at both study sites, Weldon Peete Park and the Upper Green River Biological Preserve.....	48
Figure 1.16 Boxplot comparison of the maximum call frequency of <i>Pseudacris crucifer</i> at both study sites, Weldon Peete Park and the Upper Green River Biological Preserve.....	49
Figure 2.1 Aerial photograph of the two digital recorder sites at the Upper Green River Biological Preserve in Hart County, KY used for bioacoustic monitoring in the study.....	83
Figure 2.2 Aerial photograph of the two digital recorder sites at Weldon Peete Park in Bowling Green, KY used for bioacoustic monitoring in the study.....	84
Figure 2.3 Satellite image of digital recorder locations number 1, 3, and 5 in the Munfordville, KY area.....	87
Figure 2.4 Satellite image of digital recorder locations number 2, 8, and 10 in the Greensburg, KY area.....	87
Figure 2.5 Satellite image of digital recorder locations number 11, 15, and 19 in the Upper Green River Biological Preserve, KY area.....	88
Figure 2.6 Histogram of the percentage of false positive detections by humans according to each species.....	95
Figure 2.7 Histogram of the percentage of false positive detections by human listener groups	95
Figure 2.8 Histogram of the percentage of false positive detections by human listeners for each month (April, June, October).....	96
Figure 2.9 Histogram of the percentage of computer false positive detections at Weldon Peete Park and the Upper Green River Biological Preserve.....	97
Figure 2.10 Histogram of the percentage of false positive detections by human listeners for each area with recorders (n=3).....	101
Figure 2.11 Histogram of the percentage of false positive detections by each human listener group.....	101
Figure 2.12 Histogram of the percentage of false positive detections by human listeners for each species.....	102

Figure 2.13 Histogram of the percentage of false positive detections by human listeners for each digital recorder (n=9).....	102
Figure 2.14 Histogram of the percentage of false negative detections by human listeners for each species.....	104
Figure 2.15 Histogram of the percentage of false negative detections by each of four human listener groups	104
Figure 2.16 Histogram of the percentage of false negative detections by human listeners for each digital recorder (n=9).....	105
Figure 2.17 Histogram of the percentage of computer false positive detections for each species evaluated.....	106
Figure 2.18 Histogram of the percentage of computer false positive detections for each digital recorder.....	106
Figure 2.19 Histogram of the percentage of computer false negative detections for the areas Munfordville, Greensburg, and the Upper Green River Biological Preserve.....	107
Figure 2.20 Histogram of the percentage of computer false negative detections for each digital recorder.....	108
Figure 2.21 Spectrograms of <i>Bufo americanus</i> and <i>Oecanthus</i> sp. sounds.....	115

LIST OF TABLES

Table 1.1 Information on the location of the study sites of in Weldon Peete Park and the Upper Green River Biological Preserve in south central Kentucky.....	10
Table 1.2 Summary of all anuran captures with physical methods at the Upper Green River Biological Preserve.....	28
Table 1.3 Summary of the total number of detected presences in recordings from each location at the Upper Green River Biological Preserve and Weldon Peete Park.....	30
Table 1.4 Summary of the captures of unique anurans at the Upper Green River Biological Preserve by different physical methods.....	31
Table 1.5 Summary of the anuran call characteristics in response to background amplitude at the Upper Green River Biological Preserve and Weldon Peete Park...	44
Table 1.6 Summary of linear models evaluating differences in call characteristics between the Upper Green River Biological Preserve and Weldon Peete Park.....	46
Table 1.7 Summary of the recommendations for monitoring methods for the species detected in the study.....	64
Table 1.8 Summary of the recommendations of methods based on the study design or purpose.....	65
Table 1.9 List of all herpetofaunal and mammal captures during physical surveys at the Upper Green River Biological Preserve.....	67
Table 1.10 Summary of the sound recording dates and times evaluated by Song Scope..	68
Table 2.1 Information on the location of the digital recorders at Weldon Peete Park and the Upper Green River Biological Preserve in south central Kentucky.....	83
Table 2.2 Information on the location of the digital recorder, recorder names, and coordinates of each recorder.....	86
Table 2.3 Summary of the total detections of computer and human methods in Trial I..	92
Table 2.4 Summary of the number of 3-minute recordings with the presence of anuran vocalizations in April, June, and October.....	92
Table 2.5 Results of Chi-square tests evaluating the influence of recorder location, listener group, species, and season on error rates of human and computer methods.	94
Table 2.6 Summary of the total detections of computer and human methods in Trial II..	98

Table 2.7 Results of Chi-square tests evaluating the influence of recorder area, listener group, species, and individual recorders on error rates of human and computer methods.....	100
--	-----

A COMPARISON OF CURRENT ANURAN MONITORING METHODS WITH
EMPHASIS ON THE ACCURACY OF AUTOMATIC VOCALIZATION
DETECTION SOFTWARE

Jacob Douglas Eldridge

December 2011

122 Pages

Directed by: Dr. José Pedro do Amaral, Dr. Richard G. Bowker, and Dr. Albert Meier

Department of Biology

Western Kentucky University

Currently, a variety of methods are available to monitor anurans, and little standardization of methods exists. New methods to monitor anurans have become available over the past twenty years, including PVC pipe arrays used for tree frog capture and Automated Digital Recording Systems (ADRS) used to remotely monitor calling activity. In addition to ADRS, machine-learning computer software, automated vocalization recognition software (AVRS), has been developed to automatically detect vocalizations within digital sound recordings. The use of a combination of ADRS and AVRS shows the promise to reduce the number of people, time, and resources needed for an effective call survey program. However, little research exists that uses the described tools for wildlife monitoring, especially for anuran monitoring.

In the study, there were two problems addressed relating to AVRS. The first was the poorly understood relationship between auditory survey methods and physical survey methods. I tested this problem by using current auditory monitoring methods, ADRS and the AVRS Song Scope[®] (v.3.1), alongside more traditional physical monitoring methods that included drift fences, a PVC pipe array, and visual encounter transects. No significant relationship between physical and auditory community population measures was found. Auditory methods were also effective in the detection of call characteristic

differences between urban and rural locations, further suggesting an influence of noise pollution. The second problem addressed was the call identification errors found in auditory survey methods. I examined the influence of treatments including the ADRS location, listener group, species, and season on the error rates of the AVRS Song Scope[®] (v.3.1) and groups of human listeners. Computer error rates were higher than human listeners, yet less affected by the treatments. Both studies suggested that AVRS was a viable method to monitor anuran populations, but the choice of methods should be dependent upon the species of interest and the objectives of the study.

**CHAPTER 1—A COMPARISON OF ANURAN MONITORING METHODS AND
THE EFFECTS OF NOISE POLLUTION USING A COMBINATION OF
AUTOMATIC VOCALIZATION RECOGNITION SOFTWARE AND
CONVENTIONAL FIELD SURVEY METHODS**

Abstract— Amphibian populations are declining at a rapid rate, largely due to unknown causes. Calling is an essential part of anuran reproduction, and sound pollution has been suggested as a factor in population declines. Anurans are often selectively monitored by their calls but are also regularly assessed through physical methods. Because most studies usually use only either auditory or physical surveys, little research exists comparing both methods. I hypothesized that a positive correlation between methods and observable influences of sound pollution would exist. I conducted physical surveys in two fields in the Kentucky's Green River riparian corridor using a combination of drift fences, a PVC pipe array, and visual encounter transects. Automated digital recording systems (ADRS) were also placed at the physical survey site and a similar area in a noisy city park. The automatic vocalization recognition software (AVRS) Song Scope[®] was used to search for anuran vocalizations in recordings. Comparisons of auditory and physical methods were evaluated at a community and species level. No significant relationship between physical and auditory community population measures was found, although correlation was found on a species level. Physical captures provided greater species richness but less information about some species than auditory methods. Additionally, the ADRS were effective tools in observing the effects of noise pollution as background amplitude was correlated with call characteristics. The appropriate choice of physical or auditory methods should be dependent upon the species of interest and the study's purpose, although the use of both methods can provide unique and useful information.

INTRODUCTION

Amphibians are an important part of earth's biodiversity and are often recognized as biological indicator species due to their high sensitivity to even the smallest changes in the environment. The complex life cycle of amphibians, including both aquatic and terrestrial forms, exposes them to a wider range of environmental fluctuations (Dunson et al. 1992). Furthermore, research done over the past 30 years has shown that amphibian populations are declining more rapidly and are more threatened than mammal and bird populations (Stuart et al. 2004). Many species have already become extinct, and population trend studies project that many more will soon disappear. This negative trend is further complicated by the lack of certainty about the causes of these dramatic population declines as well as the status of many amphibian populations (Stuart et al. 2004).

Numerous causes have been hypothesized for the alarming reductions in amphibian populations. The factors most frequently cited are either directly or indirectly caused by humans. Habitat loss, environmental contamination, increased levels of ultraviolet radiation, disease, climate change, human consumption and exploitation, and introduction of exotic species are often cited as having dramatic effects on species declines (Hopkins 2007). Depending on the species, combinations of these factors are likely responsible for the noted declines (Collins and Storfer 2003). Furthermore, unknown processes are involved in 48% of declining amphibian species, emphasizing the need for further research (Stuart et al. 2004). A disproportionately high amount of losses have been observed in four families of anurans (frogs and toads): Bufonidae (true toads), Ranidae (true frogs), Leptodactylidae (typical neotropical frogs), and Hylidae (tree

frogs). Habitat loss has played a fundamental role in the population declines of these four families. In particular, Ranidae has experienced population declines as a result of human exploitation as a food source, especially in Asia. Many questions still linger regarding the causes of decline in these families, especially concerning Bufonidae (Stuart et al. 2004).

In general, research concerning anurans has focused on the negative effects of chemical pollution, but recently there has been an increased concern about the effects of noise pollution on sensitive anuran species. As human populations have technologically advanced, an increasing amount of areas where anurans live and breed are disturbed by anthropogenic sound sources. These sources often include automobiles, airplanes, boats, and industry. Although the effects of these noises have been recently studied for many species of mammals and birds, studies involving anurans have been very limited (Sun and Narins 2005).

Studies of the effects of noise pollution on anuran communication have shown multiple responses by different species. Traffic noise has been shown to increase the length of time required for a female *Hyla chrysoscelis* (Cope's gray treefrog) to localize male calling, resulting in significantly less success in orienting to the male calls (Bee and Swanson 2007). Anurans have also been observed calling at higher frequencies in response to traffic noise (Parris et al. 2009). A study investigating the effects of roadway noise showed that in addition to road-crossing mortality, species with few direct road mortalities still reduced the use of breeding sites near roads with high night-time traffic (Eigenbrod et al. 2009). This difference is presumed to be the effect of the road acting as a barrier to forests or due to the traffic noise as a disturbance (Eigenbrod et al. 2009).

Conversely, some anuran populations are capable of adapting to anthropogenic noise sources. Three anuran species were shown to produce calls with different characteristics in high and low traffic noise environments (Cunnington and Fahrig 2010). When individuals in the low traffic noise environment were presented with simulated high traffic noises, they changed their calling characteristics to those similar to individuals in a high traffic environment. This call plasticity suggests that at least some species of anurans can still communicate effectively in noise-polluted environments (Cunnington and Fahrig 2010). Other anuran species have been shown to increase, decrease, or change vocalizations, potentially influencing the fitness of a given species in a noise-polluted environment (Sun and Narins 2005). This discrepancy in species-dependent response could explain why some species outperform others in a noise-polluted environment (Sun and Narins 2005).

One purpose of my study was to quantify the effects of sound pollution on anuran populations. The primary source of noise pollution was a water treatment plant that provides an average of about 60,000 kiloliters of water per day to the city of Bowling Green, KY, USA. The water is drawn from the Barren River with a noisy mechanical pump located across the river from a small local park. Other urban sounds such as vehicles and foot traffic pollute this park. In contrast, the Upper Green River Biological Preserve (UGRBP) was less disturbed. Little unnatural noise is heard at this site, especially at night.

Technological advances have brought about a new monitoring technique that allows for uninterrupted and consistent monitoring of sound-producing species. Automated Recording Systems (ARS) have been used to collect data since the

development of microprocessor-based data loggers in the 1970s. Beneficial characteristics of ARS include portability, programmability, battery power, and the ability to collect data from several sensors at scheduled time intervals (Michael and Charles 1994). Whereas traditional methods have relied on researchers conducting auditory surveys in the field, ARS has provided a more time-efficient way of monitoring anurans.

Since its inception, ARS has advanced to allow for digital recording samples during long periods, increasing the likelihood of detecting rare species or species that restrict calling to a few days of the year (Dorcas et al. 2009). It has previously been suggested that Automated Digital Recording Systems (ADRS) are more effective than traditional point and transect surveys in quality and quantity of bird and amphibian calls with the additional benefits of permanent data, data collection at any time, and the potential for automated species identification (Acevedo and Villanueva-Rivera 2006). ADRS have now become commercially available for anuran monitoring. These systems allow for long deployment times due to long battery life and large memory storage. Four of these commercially available ADRS were used in this study.

Digital recordings have allowed for machine-learning methods to be used to identify anurans, but only very recently a commercial version of automated bioacoustic recognition software has become available to researchers (Hardin et al. 2009). This software allows the researcher to sift through hundreds of hours of recordings in a short period by programming the software to recognize sound characteristics of targeted calls. Although still relatively untested, automatic vocalization recognition software (AVRS) shows promise to decrease processing time and research costs and provide better

accuracy than a human listener processing recordings or conducting field surveys (Dorcas et al. 2009). In my study, a commercially available AVRS, Song Scope[®] Bioacoustics Monitoring Software (Ver. 3.1a: Wildlife Acoustics, Inc., Concord, Massachusetts, USA), was used to analyze the digital recordings.

Even with its many advantages, ARDS has not yet provided an alternative to estimating population size or individual characteristics. In an evaluation of survey methods, ARDS detected more anuran species than point-count and transect surveys, but it lacked the ability to make an accurate density or population estimate (Acevedo and Villanueva-Rivera 2006). Furthermore, a recent study cautions heavy reliance on data from Song Scope (Ver. 2.1) because it produces a number of false positive and false negative recognition of anuran calls even after careful calibration (Waddle et al. 2009). This problem also arises in manual call surveys. Multiple studies have recorded that even small error rates of human listeners lead to considerable effects on site occupancy models (Royle and Link 2006). Consequently, the benefits of AVRS should be weighed against human error by both manually counting calls from a recording and conducting field surveys (Dorcas et al. 2009). Therefore, for a large quantity of calls and limited human involvement, AVRS could be a more effective method.

Traditional methods such as drift fences, pitfalls, funnels, crayfish traps, transect surveys, visual encounter surveys, egg counts, tadpole counts, cover objects, PVC pipe arrays, larval litter bags, sound recordings, and net sweeps are all used to measure herpetofauna (Dodd Jr. 2003). Although many of these methods are often used in conjunction with one another, little is known about the connection between these traditional survey methods and ARDS. A comparison of intensive physical sampling,

standardized call surveys, and ARS with only manual listening suggested that standardized call surveys may be the most cost-effective and efficient choice for a trained professional (Muths and Iko 2000). In a previous comparison of the calling rates to mark-recapture population estimates of *Lithobates clamitans* (green frog), a positive relationship was seen between both methods (Nelson and Graves 2006). Another study showed that ADRS detected more bird and anuran species than traditional methods including fixed-radius point counts and transects (Acevedo and Villanueva-Rivera 2006). Alternatively, other studies caution reliance on call surveys due to the high degree of human error both from field surveys (Lotz and Allen 2007) and from digital recordings (Genet and Sargent 2003).

The objective of this study was to quantitatively compare the density of calling anurans from ADRS to the number of anurans captured by traditional field survey methods. The relationship between methods highlights the advantages and disadvantages of each, providing a suggestion for an optimum monitoring method for anurans. An additional goal was to examine the effects of sound pollution on the anurans of Kentucky. I hypothesized that there would be a positive correlation between auditory and physical survey methods. Moreover, I predicted that there would be a response of anuran call characteristics in response to background amplitude, and call characteristics would be different between an urban and rural environment.

METHODS

Study Sites—I conducted my research at two sites. The first site was a local city park in Warren County, Weldon Peete Park (WPP), located within the Bowling Green, KY, USA city limits. This location was used as an example of a sound-polluted location.

The other site was located at the Upper Green River Biological Preserve (UGRBP), managed by Western Kentucky University. This site was an example of a natural area that is unaffected by sound pollution. A more detailed overview of each location, which includes a general vegetation survey may be found elsewhere (Trimboli 2010).

WPP is located within the city limits of Bowling Green, Kentucky. Bowling Green is the third largest city in Kentucky with 58,067 residents in a land area of 97.85 km² in the 2010 census (Bureau, U.S. Census. 2011. *Kentucky State and County Quick Facts*. Available from census.gov [accessed 11 November 2011]). WPP (also known as Mitch McConnell Park) is approximately 30 ha of deciduous forests mixed with hayed fields, and a paved bike and walking path runs throughout its entirety (Trimboli 2010). The park is bordered by a main road and the Barren River, a tributary of the Green River. Nearby, there is a large hospital (0.5 km) and a railroad (0.63 km) (Trimboli 2010). Song Meters were termed BG01 and BG02 (Table 1.1; Figure 1.1). BG01 was located on the edge of a wooded area and field edge approximately 50 m from the Green River. There is a temporary pool present through much of the spring near BG01. BG02 was approximately 300m away from BG01 and was positioned about 10m away from the river in the wooded river corridor. BG02 was closer to the main road and the Bowling Green Municipal Utilities Water Treatment Plant.

The UGRBP was purchased and designated as a preserve in 2004 and is used for education and research. It is approximately 475 ha and encompasses a variety of habitats along the Green River corridor and largely consists of deciduous forest and old fields, some of which are being restored to their natural state. Access to the preserve is by permission only. The farthest extent of the UGRBP border is approximately 2 km from

Mammoth Cave National Park. Areas surrounding the UGRBP are primarily farmland, and the closest town is Horse Cave, KY approximately 5 km away in Hart County. In 2010, Hart County had a population of 18,199 (Bureau, U.S. Census. 2011. *Kentucky State and County Quick Facts*. Available from census.gov [accessed 11 November 2011]).

Two ADRS were placed and physical monitoring was conducted in two bordering fields at the UGRBP. Field A contained the recorder site BP01, and field B contained the recorder site BPO2 (Table 1.1; Figure 1.2). Both recorders were placed in a wooded area within the riparian corridor between the field and the Green River. Both fields were completely surrounded by at least 10 m of wooded area but generally more than 50 m. Field A is being restored to bottomland hardwoods using tree planting. It was planted lengthwise with saplings (about 1–2.5-m tall) in rows separated every 4 m by underbrush that is occasionally mowed between the rows. Field B was used as a hay field and was hayed by a farmer three times during this study. Both fields were subject to infrequent flooding of the Green River during exceptionally high precipitation events. Pooling water was common in the spring. However, standing water remained longer in field A in both rutted areas and an area where a temporary pool remained at least throughout the spring and occasionally later in the year.

TABLE 1.1 Summary of the location, recorder name, and coordinates of recorders used in the study. Weldon Peete Park is located in Bowling Green, KY and has the recorders BG at the site; the Upper Green River Biological Preserve is located in Hart County, KY and BP are located at the site

Location	Recorder	Coordinates
Weldon Peete Park	BG01	N37° 00' 0.003" W86° 25' 24.6"
	BG02	N36° 59' 58.2" W86° 25' 33.00"
Upper Green River Biological Preserve	BP01	N37° 14' 57.3" W85° 59' 29.5"
	BP02	N37° 14' 53.9" W85° 59' 13.8"



FIGURE 1.1 Aerial photograph of the recorder sites (BG01 and BG02) at Weldon Peete Park in the city of Bowling Green, KY used for auditory monitoring

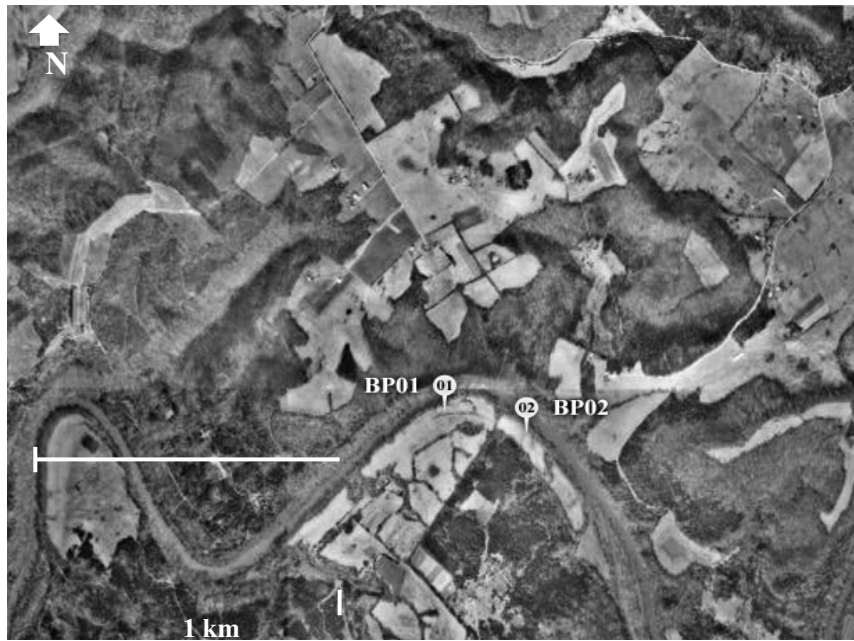


FIGURE 1.2 Aerial photograph of the recorder sites (BP01 and BP02) at the Upper Green River Biological Preserve in Hart County, KY used for bioacoustic monitoring.

Comparison of Methods—Six weatherproof Song Meter digital recorders (Model SM2, Wildlife Acoustics, Inc., Concord, MA, USA) were used in this study. Two ADRS were located approximately 100m apart at WPP. Four ADRS were located at the UGRBP with duplicate recorders approximately 300m apart, each at the edge habitat of separate fields. These systems use two microphones per recorder. Settings were adjusted such that one microphone gain was 3dB more than the other on the same recorder. Each microphone was covered with a foam windscreen to reduce wind noise interference. The ADRS were set to record the first three minutes of each hour in a WAV format on one of two memory cards in each recording device. After the recorders were damaged by unidentified mammals, I built protective cages to surround the recorders using galvanized steel hardware cloth. I placed two HOBO[®] dataloggers at the same location of an ADRS at both the UGRBP and WPP. A HOBO[®] U23 Pro v2 Temperature and Humidity

datalogger recorded both temperature and humidity at 15-minute intervals. A HOBO[®] UA-002-64 Temperature and Light pendant datalogger recorded temperature and light intensity at 15 minute intervals.

To evaluate the sound recordings, I used the AVRS Song Scope[®] Bioacoustics Monitoring Software (ver. 3.1a: Wildlife Acoustics, Inc., Concord, Massachusetts, USA) to evaluate the data collected by the ARDS. I used clear, identifiable calls to build recognizer files for anuran species in Song Scope. Song Scope recognizer files are a compilation of characteristics from annotated vocalizations that are used for comparison to similar sounds from recordings. All vocalizations used in the construction of recognizer programs were either from the UGRBP or areas within 50 km to avoid local differences in calls. I used the software manual's guidelines to adjust parameters (e.g. dynamic range, call length, Fast-Fourier Transform size, etc.) used to create recognizer files. I created recognizer files for the eight species that had adequate vocalizations in previous recordings or my current study sound data. These species included *Bufo americanus* (American toad), *Bufo fowleri* (Fowler's toad), *Pseudacris crucifer* (spring peeper), *Pseudacris feriarum* (upland chorus frog), *Lithobates clamitans* (green frog), *Lithobates catesbeiana* (bullfrog), *Hyla chrysoscelis* (Cope's gray treefrog), and *Acris crepitans* (northern cricket frog).

All recognizer files were then tested for error rates by comparing AVRS-detected vocalizations and human-detected vocalizations by manual listening. A recognizer score is an adjustable parameter in Song Scope that represents the "fit" of a candidate sound to the recognizer model. Depending upon the error rates of a recognizer, I reduced the minimum score of the recognizer model to decrease false negatives. Scores were reduced

only to the extent that would decrease false negatives without greatly increasing false positives. I then used the recognizer files for each species to search available audio recordings during the study period for candidate vocalizations (those resembling model vocalizations). I validated each candidate vocalization, recording the presence or absence of calls and the number of correct detections during each hourly sample.

The UGRBP served as the study site for traditional monitoring techniques (Figure 1.3). The physical methods were designed to capture the anuran species that have been identified at Mammoth Cave National Park, located 4 km downstream of the study site, and species previously identified at the UGRBP. The area with a temporary pool was marked by a change in vegetation, and I recorded an estimate of the water diameter on each trap day (Figure 1.3). Four drift fence arrays, 32 polyvinyl chloride (PVC) pipes, and eight transects were used each trap day to survey the UGRBP herpetofauna.

Drift fences are often used in herpetofaunal surveys to measure a wide variety of species and have been shown to be more effective than visual encounter surveys for comprehensive population sampling (Crosswhite et al. 1999). Drift fence arrays in combination with a mark-recapture analysis of individuals can be used to make an accurate estimate of a population (Halliday 1996). They are often composed of both pitfall and funnel traps. However, only large pitfall traps were used because they more effectively capture anuran species (Greenburg et al. 1994). Additionally, increased maintenance and mortality, especially in anuran species, may be associated with funnel traps (Enge 2001).

The drift fences were located in two bordering fields adjacent to the Green River (Figure 1.3). Areas immediately adjacent to the drift fences were composed of similar

habitat. There were two drift fences erected on the field-woodland edge habitat in each field. Each drift fence was aligned parallel to the Green River. In each field, the drift fence nearest the river was located approximately 15 m from the Green River and within 60 m of the ADRS in that field. The other fence was located approximately 80 m across the field along the opposite field-woodland edge habitat (Figure 1.3). Each drift fence array was constructed in a t-shape with three 10 m arms radiating from a central point. The “top” of the t-shape (approx. 20 m) paralleled the edge habitat between the field and woods, and the other arm intersected “top” and ran 10 m into the wooded habitat.

The drift fence was constructed using 0.91 m high (4 ft.) silt fence stapled to wooden stakes. A thin trench was dug using a shovel to outline where the fence would be positioned. The silt fence was buried in the trench approximately 10-cm deep. An auger mounted on a tractor was used to dig seven large holes for pitfall traps at each drift fence location. One pitfall trap was located at the intersection of the fence arms. A pitfall trap was also positioned at both 5 m and 10 m on each arm of the drift fence. Pitfall traps were made of 19 L (5 gal.) buckets with lids. The lids from the buckets were cut making an approximately 5-cm wide lip to prevent the escape of captured animals. Pitfall traps were placed with the opening flush or slightly below ground level. The silt fence was placed to intersect the center of each pitfall opening. Wet sponges were placed at the bottom of each trap when the trap was open to prevent desiccation or drowning of the trapped animals (Todd, et al. 2007). When traps were not in use, a cover made of aluminum flashing and the removed center of the bucket lid was placed securely over the pitfall opening.

Although drift fences have proven effective in trapping bufonids and ranids, they are less effective for trapping hylids because they may more easily climb out and escape from the pitfall and funnel traps. To better capture hylids, I used PVC pipe arrays in conjunction with the drift fences. PVC pipe arrays have been used successfully to capture multiple Hylidae species (Boughton and Staiger 2000). A hylid anuran known to exist at the study site, *Hyla chrysoscelis*, Cope's gray treefrog, has been successfully captured using PVC pipe arrays (Pittman et al. 2008).

A PVC pipe array was organized in each field at the UGRBP with an ADRS serving as a central point (Figure 1.3). There were 32 white PVC pipes that followed the size specifications of Pittman for attracting *Hyla chrysoscelis* (Pittman et al. 2008). Pipes had an inside opening of approximately 3.8 cm and were approximately 1.5 m tall. Each pipe was driven into the ground approximately 15 cm. No water was intentionally placed into the pipes, but small amounts of water occasionally accumulated in the bottom of the pipes after a rain event. Pipes were positioned relative to the ADRS. Eight pipes were spaced 25m apart along the edge habitat on the field and woods with the ADRS centered between the two centermost pipes (Figure 1.3). Pipes were positioned approximately 5m from the field edge. Another 8 poles were placed directly across from these poles on the opposite field edge (Figure 1.3). This layout of 16 poles was replicated in both fields (Figure 1.3).

I also incorporated visual encounter surveys by walking multiple transects. Visual encounter surveys are known to be effective in long term monitoring of amphibians, especially near breeding sites, and may be less biased in the species captured than the other physical survey methods (Crump and Scott 1994). I walked four

transects in each field on dates when I checked pitfall traps and the PVC pipe array (figure 1.3). All transects were parallel to the PVC pipe rows. In field A, the width of each transect was the 4 m between each tree row, allowing for 19 transects (Figure 1.3). There were fewer transects in Field B ($n=12$) due to the narrowing of the field in the center (Figure 1.3).

I used a random number generator to designate the two transects that I would walk in each field on a given day. In addition, I always walked two of the four transects in each field along the field edges bordering the PVC pipe rows (Figure 1.3). The PVC pipes were checked when walking these transects, and the short distance to and from the PVC pipe was also counted as part of the edge transects. Along the field edge furthest from the river, the transect was a two-track used occasionally for preserve maintenance and research (Figure 1.3). All transects began at the line between the first PVC pipes on opposite sides of the field and ended at the line between the last two PVC pipes on opposite sides of the field. Therefore, the total length of each transect within the field is approximately 150 m, and each transect walked along the edge habitat was this length plus an additional approximately 40 m that was walked to check the PVC pipes. I alternated daily the field where transect surveys were completed first. To survey for anurans in these transects, I walked at a normal pace and scanned the ground in front of me. If an anuran was observed, I attempted to capture it and process it, taking note of its location. If the anuran was not captured, the species (if discernible) and location was noted although it was not marked. These eight transects were completed prior to monitoring the drift fence pitfalls.

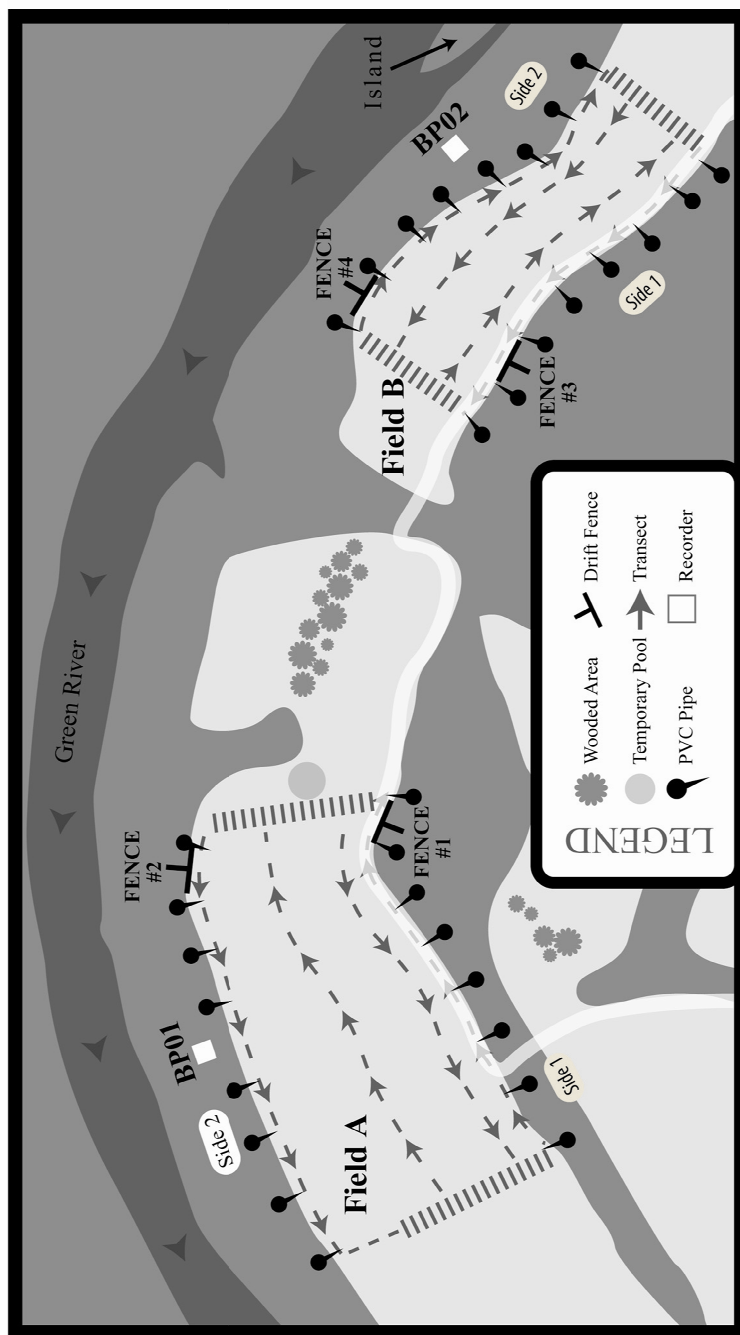


FIGURE 1.3 A schematic of the study site at the Upper Green River Biological Preserve in Hart County, KY. Markers are representative of the study method performed on each trap day of the study. The temporary pool was present approximately March–June. Wooded areas include all those of the described shad. Fence #1–4 are drift fences with 7 pitfall traps each. The island is above water during normal river flow.

Captured anurans were marked by cutting the end of a single toe (toe-clipping) and photographing individuals. The positive and negative aspects of effective and ethical individual marking have been highly debated. Although studies have suggested that toe-clipping may have negative effects on recapture rates of amphibians (McCarthy and Parris 2004), there is no conclusive evidence that these statistical estimations are a result of toe clipping (Philiott, et al. 2007). In fact, toe clipping may be the least stressful method of marking amphibians when done properly (Philiott, et al. 2007). To reduce the potential adverse effects of toe clipping, I only clipped one toe per individual. This toe was the digit one, three, four, or five on the rear foot to reduce potential harm. I marked captured anuran individuals by either clipping a toe at a joint for non-Hylidae species or by removing a toe pad for hylids at the time of the first capture. Toe-clipping followed the standard operating procedures as described by the USGS National Wildlife Health Center (Green, D. E. 2001. Toe clipping of frogs and toads. USGS, available from nwhc.usgs.gov [accessed 18 February 2010]). All physical capture and marking techniques were done in accordance with Western Kentucky University Institutional Animal Care and Use Committee (Animal Welfare Assurance #A3558-01, protocol 10-03).

Individuals were then photographed to aid in further identification by comparing the skin markings. If an individual was recaptured with a previously clipped toe, it was photographed and identified according to the toe clip and pattern variation. I then released animals that were caught in pitfall traps at least 15 m away from the fence after processing. Animals inhabiting the PVC pipes were returned back into the PVC pipe where they were found. Anurans that were identified to a species level but were not

captured, were dead, or hatched within the same year were also recorded but not processed. I began additional processing steps beginning in October 2010. This included placing the anurans in a modified squeeze box (Cross 2000). I then placed a clear transparency over the individual and made lines for the snout and vent locations. The length between these marks were measured at a later time with a caliper and recorded as the snout-vent length. I also measured mass by placing the individual in a Ziploc™ bag and attaching it to a calibrated Pesola™ spring scale. Mass was measured to the nearest gram. Pitfalls were closed when not being used during a trapping period. I did not process anurans daily if continually captured due to their residence in a PVC pipe and their pattern and the toe-clip mark was easily recognized, however, habitual pipe residents were completely processed at least once every two weeks.

Field research was conducted during four periods emphasizing the seasons of greatest anuran activity, spring and summer. PVC pipe arrays were set on 4/02/2010. Drift fence arrays were constructed from 31 May–4 June 2010 in the first field season. The first summer season of data collection occurred from 6 June–30 July 2010. Two ADRS were deployed on 15 July 2010. On 30 July 2010, the pitfalls were closed and covered with a pile of dirt to ensure their long-term closure. The second season of data collection occurred during the dry season and lasted from 18 October–29 October 2010. Two ADRS were also used during this period. Drift fences were removed on 29 October 2010 and pitfalls were covered by piles of dirt and marked with flags. New drift fences were constructed from 30 March–2 April 2011. The third period encompassed the spring and summer season ranging from 4 April–29 July 2011. In this trap season, two replicate ADRS units were used at each recording site for a total of four recorders at the UGRBP.

Recorders were serviced in pairs (one from each field monthly) on alternating schedules to ensure a complete auditory data set in case of a recorder malfunction. Two ADRS were also deployed at both WPP recording sites during the 2011 season. During April 2011, two weeks were represented by only one trap day because traps were closed and recorders were removed from each site when river flooding was forecasted to overtake the study areas. The dates when traps were closed and recordings were stopped were April 13–15 and April 27–29 except for recorders at Weldon Peete Park that were not removed during April 13-15. After traps were closed, all equipment was removed from the field besides the ADRS and environmental dataloggers.

During these periods, traps were opened on Monday morning between 0800 and 1000 CST. There were no anurans marked or transect surveys conducted on Mondays, but chance visual observations were noted. Tuesdays through Fridays, I completed transects and checked traps approximately every twenty-four hours. I began surveys between 0800 and 0900 and ended between 0900 and 1130. Occasionally, I would need to bail water out of the traps after heavy rain events. I recorded when pitfall traps were completely full of water, potentially allowing captured individuals to escape.

Herpetofauna outside of transects or traps but within the study area were also recorded but never marked. Small mammals in traps were also recorded but not identified to the species level. Suspected influence of larger animals tampering with traps (e.g. sponges moved with no animals present) was noticed beginning on 8 June 2011. Therefore, three live mammal traps were set and baited beginning on 23 June 2011, and captured animals were relocated approximately 3 km away from the study site.

Sound Pollution—To test the effects of sound pollution, recordings from WPP and the UGRBP were analyzed using the software Raven (Raven Version 1.2.1 2003-2004). Sound analysis was conducted only on the recorder with the most vocalizations detected using AVRS from each WPP (BG01) and the UGRBP (BP01). Only species with ten or more recordings with detected presences at both sites were used in the analysis. Those species were *Bufo americanus*, *Hyla chrysoscelis*, and *Pseudacris crucifer*. Five recordings were randomly chosen from those with the species of interest present using a random number generator. From the total of 180 seconds of sound on each recording, I used a random number generator to select three times within each evaluated track. I then selected the closest distinct vocalization nearest to the randomly selected time. If a single vocalization was nearest to multiple time selections, then the next closest vocalization was measured. To reduce the effects of unwanted environmental parameters, I selected recordings from the UGRBP to correspond with the date and time as the WPP recording. If the species was not calling at that date and time, another recording was chosen that most closely reflected the date, time, and temperature of the recording from WPP. A spectrogram constructed in Raven was used to measure the vocalization frequency maximum and minimum in hertz (Hz) as well as its length in seconds (s). Additional measurements included the maximum amplitude and root mean square (RMS) amplitude during the time of the call. Amplitude measurements were relative to the settings of the recorder. All recorders and their respective settings were identical. A total of 15 vocalizations from each site and for each species were analyzed using these methods.

Statistical methods—I used SPSS statistical software 19.0.0 (PASW Statistics 2009) and R statistical software (R Development Core Team) for data analysis. Acceptable probability values were set at $\alpha=0.05$ unless otherwise noted. To reduce bias, auditory and physical methods were conducted by a single researcher (myself). I termed anurans that hatched during the year of capture as metamorphs, and they were not included in the majority of data analyses. Adults were considered to be individuals that were clearly larger than metamorphs. Individuals were seldom sexed, therefore sex was not included as a factor in the study. Anuran captures were considered independent because individuals were identified. Adult individuals that were not captured but were identified to species level were called unmarked. Unmarked anurans were included in tests regarding density comparisons along with recaptures.

Auditory data from recorders at the same study site (e.g. the two recorders at the UGRBP or WPP approximately 300-m apart) were evaluated for the detection of the same calls by comparing tracks from the same time for all species. Recorders were determined to be in sync with one another, based on loud noises at the study site (e.g. gunshots). Furthermore, no apparent duplication of anuran vocalizations in recordings at the same time was observed between the two recorders. Therefore data from each recorder were treated independently from other recorders. Auditory data from recorders within the same site, either the UGRBP or WPP, were summed for certain analyses.

Although the number of vocalizations detected in each recording was documented, only the presence and absence of species during a given recording was considered in all analyses. The reasons for this were twofold: first, the number of calls may not directly correlate with the number of individuals (e.g. one individual could call

ten times during a recording or ten individuals could each vocalize once in a recording). Second, the number of vocalizations detected by the recognizer may not accurately reflect the number of calls actually occurring (e.g. the blending of calls in a strong chorus may be difficult for the recognizer to individually discern and count the same number of calls as only a few calls occurring in quieter conditions).

The number of recording with the presence of a species during a trap day (maximum 24) was used as a measure of call density to compare to physical capture methods. A trap day was considered from 1000 hrs from the previous day to 0900 hrs of the trap date. These times were chosen to reflect the average time that traps were checked the previous day through the time before traps were checked the next day. Therefore, physical captures of the morning of 25 May 2011 would be compared to auditory data from 24 May 2011 at 1000 to 25 May 2011 at 0900. This auditory time sequence would likely more accurately reproduce the activity detected by physical trapping methods.

Physical methods comparison—The physical methods used in this study (drift fences, PVC pipe array, and transects) were compared based on their ability to capture new adult individuals. The total number of adult individuals captured by each method was assessed using a Chi-square test.

Species richness comparisons—The effectiveness of auditory and physical methods to determine species richness on a weekly and daily scale was also evaluated. The species richness detected by auditory methods during each week and day was compared to the species richness determined by physical methods during each week and day using a Mann-Whitney U test.

Community comparisons—An overall comparison of the physical and auditory methods throughout time was tested by a comparison of dissimilarity matrices using Mantel tests. Weeks were the factors of matrix rows, and the species were the variables of matrix columns. The Bray-Curtis dissimilarity index was chosen to compute the dissimilarity matrix using the R package *vegan* (Oksanen et al. 2011). The Sorensen dissimilarity index was used to compare presence-absence data (Legendre and Legendre 1998). This ecological distance semi-metric was chosen because it only compares similarities of only present values, eliminating the potential inflation in similarity due to double-zeros. Double-presence of zeros is often considered by ecologists to be a poor measure of actual similarity, and conclusions from these measures are generally discouraged (Legendre and Legendre 1998). The Bray-Curtis dissimilarity is calculated using the double weight of similarity divided by the minimum abundance at each location (or week in this case). Therefore, species with no presences in each matrix as well as weeks with no presences were removed from each matrix. The species removed were *L. clamitans* and *P. feriarum*, and removed rows were the two weeks from October 2010 and the last three weeks of July 2011. The dissimilarity in the composition of species ($n=6$) during each week ($n=16$) was compared using a Mantel test with 10,000 permutations. Both trap presence and absence data and weekly capture means were compared to auditory presence-absence and mean weekly call detection presences. Furthermore, the error rates that were determined for each recognizer model were used to adjust the auditory data for an additional comparison. The Mantel tests examined the hypothesis that the degree of dissimilarity between sampling methods corresponds to the degree of dissimilarity in anuran species compositions.

Species-level comparisons—Further comparisons of physical and auditory detections on a species-by-species scale were conducted using Spearman correlations and Chi-square tests. Pearson's Chi-square randomization tests (due to low numbers) with 10,000 permutations were used compare the presence of captures during weeks when a given species calling to weeks when it was not calling. Enough data by weekly detections were only present for *Bufo americanus*, *Bufo fowleri*, *Hyla chrysoscelis*, and *Pseudacris crucifer*. Because *H. chrysoscelis* was captured in all weeks, it could not be statistically evaluated in a meaningful way. Because I also was interested in the correlation of movement during the calling season, data were reduced to the first period (either day or week) to the last period when the species was detected calling. All physical detections of adult individuals were included (captures, recaptures, and non-captures). Correlations were tested by comparing the daily and weekly sum of recordings or mean weekly presences during trap nights to the number of physical captures. The only two species with sufficient data for meaningful statistical comparisons during calling seasons were *B. fowleri* and *H. chrysoscelis*. Data were based on frequencies, therefore a Spearman correlation was used to test significance of associations.

Time of day comparisons— A series of exploratory analyses were conducted to detect potential effects of sound pollution. To test whether sound pollution influenced the calling activity during different periods of the day, the temporal calling trends between WPP and the UGRBP were examined. Only data prior to 23 June 2011 were used for the comparisons due to the recording absence after this time at WPP. The time of calling was grouped by the number of call presences detected during six given periods of the day that were based upon sunrise and sunset. The times of the sunrise in the study

varied from 0523–0629 hrs, and sunset varied from 1906–2009 hrs. These periods were: period 1= 0200–0500 hrs (night pre-sunrise), period 2= 0600–0900 hrs (morning sunrise), period 3= 1000–1300 hrs (midday), period 4= 1400–1700 hrs (early evening), period 5= 1800–2100 hrs (evening sunset), period 6= 2200–0100 hrs (night post-sunset). A correlation between periods was performed to measure the degree of association. Because data were frequencies, Spearman's correlation was used to compare rank associations.

Calling season comparisons—To examine the differences in calling seasons at WPP and the UGRBP, I compared the days when a species was calling. I visually compared the calling activity of each of the three species by dates with detected calls in a recording.

Influences of background amplitude—To test the potential effect of background noise on call characteristics, results from sound analysis from WPP and the UGRBP were examined. Potential call plasticity due to background amplitude was tested using a series of regressions that compared the amplitude during the time of the vocalization to both the maximum frequency (Hz) and length (s) of the call for each species at each study site.

Comparison of call characteristics by location—To test if the call characteristics at an urban (WPP) and rural (UGRBP) location were different, call characteristics from each location were compared to each other. A general linear model was constructed for each species that compared location (fixed factor) to response variables including temperature, length of call, maximum frequency of call, and maximum and RMS amplitude during the time of the call.

RESULTS

Twelve of the 13 species observed during this study at the UGRBP were detected using physical survey methods (Table 1.2). From a total of 102 trap nights, 147 live, adult anurans representing ten species were captured and identified using physical survey methods. An additional 19 more adult anurans were observed but were not marked because they were dead or evaded capture. Of these individuals that were unmarked, the only *Gastrophryne carolinensis* (Eastern narrowmouth toad) observed in this study was found dead in a flooded pitfall trap. A total of 301 recaptures of adult anurans were recorded with majority of recaptures (n=290) being of *Hyla chrysoscelis* (Cope's gray tree frog). 231 metamorph anurans (hatched during year of capture) representing seven species were identified to species level (table 1.2). *Lithobates sylvatica* (wood frog) was only observed as metamorphs. Additional toad metamorphs (n=103) from June–July 2010 were not distinguished as either *Bufo americanus* or *Bufo fowleri*.

TABLE 1.2 Summary of anuran captures using physical methods at the Upper Green River Biological Preserve. Metamorph=hatched during the year of capture, Unmarked= not captured or dead

Species	Unique adult first captures	Adults Unmarked ²	Number of recaptures	Metamorph captures ³
<i>Acris crepitans</i> (Northern cricket frog)	13	0	0	0
<i>Bufo americanus</i> (American toad)	11	0	6	25
<i>Bufo fowleri</i> (Fowler's toad)	39	0	0	50
<i>Gastrophryne carolinensis</i> (Eastern narrowmouth toad)	0	1	0	0
<i>Hyla chrysoscelis</i> (Cope's gray treefrog)	46	2	290	1
<i>Lithobates catesbeiana</i> (bullfrog)	2	0	0	0
<i>Lithobates clamitans</i> (green frog)	4	0	0	3
<i>Lithobates palustris</i> (pickerel frog)	18	13	0	0
<i>Lithobates sphenoccephala</i> (Southern leopard frog)	5	2	0	88
<i>Pseudacris crucifer</i> (spring peeper)	7	1	5	54
<i>Pseudacris feriarum</i> ¹ (upland chorus frog)	0	0	0	0
<i>Lithobates sylvatica</i> (wood frog)	0	0	0	10
<i>Scaphiopus holbrookii</i> (Eastern spadefoot)	2	0	0	0
¹ No individuals were physically observed but were detected with audio surveys ² Unmarked adults were either dead or were not captured in study area and could potentially be recaptures ³ Metamorph toads were not distinguished between Fowler's and American toads in June–July 2010 (109 total) and are not represented on table				

Audio recordings were collected from one ADRS at the UGRBP from 15 July 2010 at 1100 to 30 July 2010 at 2300. The other ADRS only recorded from 15 July

2010 at 1100 to 27 July 2010 at 2000. Both ADRS at the UGRBP successfully recorded during the October trap period. Recordings were also successfully retrieved from ADRS during the entire 2011 trap season at the UGRBP except during the periods when units were removed due to flooding concerns. For an unknown reason, the recorders did not function at Weldon Peete Park after 23 June 2011. All recordings were successfully retrieved at this location before this time besides during flooding periods in April (Appendix 1). No significant malfunctions rendering any recordings unusable were noted.

Sound recordings from the ADRS units were analyzed during each period for the eight species with recognizer models using AVRS (Table 1.3). A total of 10,169 recordings representing 30,507 minutes of sound were evaluated. The mean time of a recognizer to evaluate 100 recordings (300 min) was 9.4 min. Therefore, sound was processed at a mean rate of 32 minutes of sound data per one minute of real time. There was a total of 45,610 detections for all eight species and a mean number of 0.66 results per recording. Of these detections, 9,145 were validated as anuran vocalizations, resulting in an overall true positive detection rate approximately 20%. The AVRS, Song Scope, processed recordings for a total of approximately 127 hrs. The time required to validate 1000 AVRS results was approximately 30 min. but varied according to the number of actual presences that were recorded. Therefore, the total time required to validate results was about 23 hours.

During trap days when recorders and physical methods were in use simultaneously, all the sound recordings were analyzed for comparison with physical

capture data. These data sets were of eight days (19.2 hours) in July 2010, eight days (19.2 hours) in October 2010, and 62 days (148.8 hours) in April–July 2011.

The presences of anuran calls in recordings varied highly by species (Table 1.3). The independence of recorders was confirmed as the two recorders at the same study site (either UGRBP or WPP) recorded the same species at the same time in only 73 instances (1.9%) of 3,789 possible opportunities of replication. Of the 1,682 hourly recordings during which recorders at the four locations were functional (excludes two flooding periods in April and after June 23 at 1400 hrs, 505 recordings (30.0%) had at least one species present.

TABLE 1.3 The number of detected recording (3min) presences of each species at recording locations per species at the Upper Green River Biological Preserve and Weldon Peete Park.

Species	BP01 (UGRBP field A)	BP02 (UGRBP field B)	BG01 (WPP site 1) ¹	BG02 (WPP site 2) ¹
<i>Acris crepitans</i>	3	1	0	0
<i>Bufo americanus</i>	120	35	10	0
<i>Bufo fowleri</i>	43	141	0	0
<i>Hyla chrysoscelis</i>	158	154	35	3
<i>Lithobates catesbeiana</i>	2	0	0	0
<i>Lithobates clamitans</i>	1	0	1	0
<i>Pseudacris crucifer</i>	216	11	15	3
<i>Pseudacris feriarum</i>	9	2	0	0
¹ BG01 and BG02 are limited to times on or before 6/23/2011 at 1400 hours				

Physical methods comparison—The observed frequency of the original capture of anurans according the three physical capture methods (drift fences, PVC pipes, and transects) was significantly different than the expected distribution ($\chi^2=26.663$, $df=2$, $P<0.0001$). Drift fences had the most original captures ($n=72$), followed by the PVC pipe

array (n=51) and transects (n=24). Regarding species richness using each method, drift fences also had the highest number of live, adult anuran species (n=9), followed by transects (n=7), and PVC pipes (n=2) (Table 1.4). PVC pipes provided the majority of captures for the arboreal species in this study, *Hyla chrysoscelis* and *Pseudacris crucifer*. Captures were recorded in 22 of 32 pipes (69%) during the study. The pipe with the most unique adult individuals captured during the course of the study was pipe 7 (n=7) followed by pipe 8 (n=6). Pipe 8 captured the most individuals at a single time (n=3). Only two of the individuals captured in PVC pipes were recaptured in a different pipe than the pipe originally occupied. One *H. chrysoscelis* and one *P. crucifer* individual were captured in transects, and one *H. chrysoscelis* individual was captured against a drift fence. Otherwise, terrestrial species were captured more often by drift fences than transect visual encounter surveys except for *Lithobates sphenoccephala* (Southern leopard frog), most often by transects (table 1.4).

TABLE 1.4 Summary of captures of unique anurans at the Upper Green River Biological preserve during the course of the study.

Species	Drift Fence	PVC Pipe	Transect
<i>Acris crepitans</i>	7	0	6
<i>Bufo americanus</i>	8	0	3
<i>Bufo fowleri</i>	30	0	9
<i>Hyla chrysoscelis</i>	1	44	1
<i>Lithobates catesbeiana</i>	2	0	3
<i>Lithobates clamitans</i>	4	0	0
<i>Litobates palustris</i>	16	0	2
<i>Lithobates sphenoccephala</i>	2	0	3
<i>Pseudacris crucifer</i>	0	6	1
<i>Scaphiopus holbrookii</i>	2	0	0

The mean daily species richness for auditory methods, physical adult captures, physical adult plus hatchling captures, and total detection using both methods were 1.09 (sd=0.983), 1.68 (sd=0.798), 2.56 (sd=1.465), and 3.05 (sd=1.502), respectively. Ranked daily auditory species richness was found to be significantly different than physical methods considering adults only ($U=2003.5$, $n=78$, $P=0.0001$) and adults plus hatchlings ($U=1285$, $n=78$, $P<0.0001$).

Species richness comparisons—The mean weekly species richness for auditory methods, physical adult captures, physical adult plus hatchling captures, and total detection using both methods were 1.71 (sd=1.146), 2.95 (sd=1.431), 3.95 (sd=1.987), and 4.86 (sd=1.852), respectively. Ranked weekly auditory species richness was found to be significantly less than physical methods considering adults only ($W=117$, $n=21$, $P=0.007$) and adults plus hatchlings ($W=78$, $n=21$, $P=0.0003$).

Monthly species richness was visually compared in 2011 (Figure 1.4). The month of June had the highest amount of total species richness with nine species observed. The species detected in 2011 that were not detected during June was *Acris crepitans* and *Pseudacris feriarum*, both limited to detections in April and May 2011. Detection of adult anurans using physical methods was variable throughout each month. However, species richness of physical captures with hatchlings was higher in June and July due to the presence of a greater diversity of hatchlings in these months.

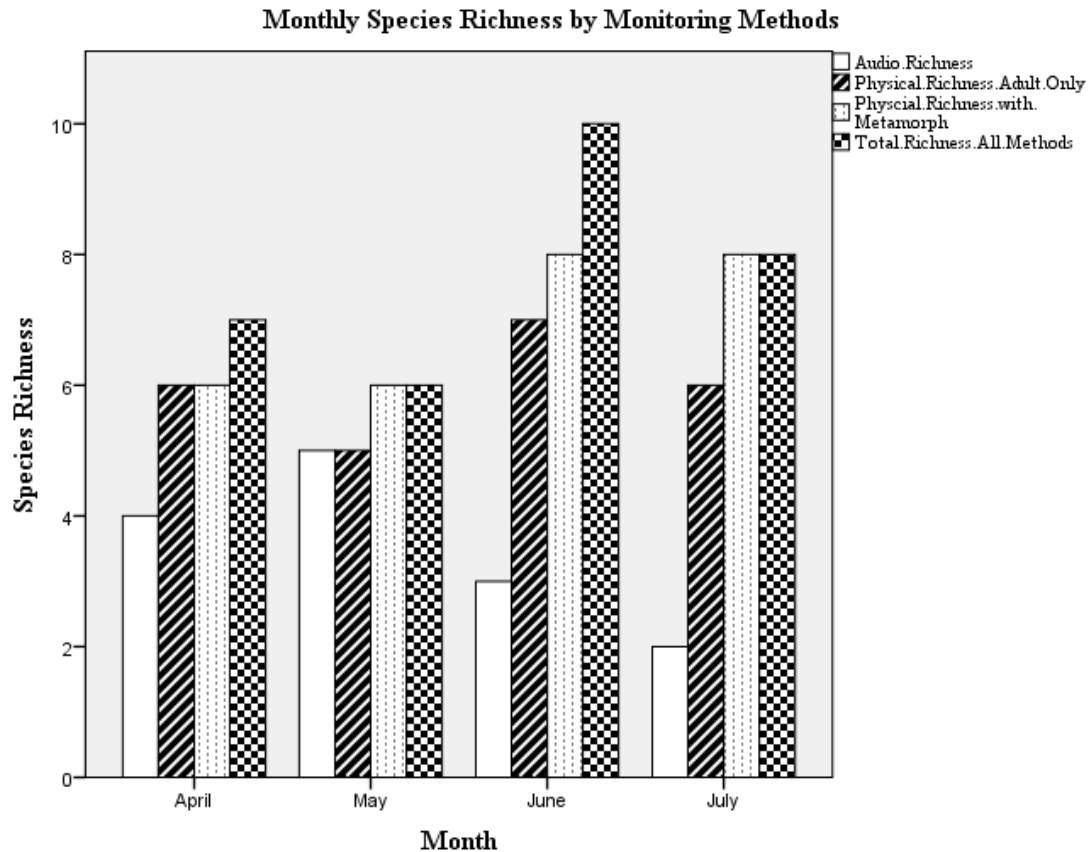


FIGURE 1.4 A comparison of the total species richness detected by methods employed in this study at the Upper Green River Biological Preserve during the months of monitoring in 2011. Audio.richness=digital recording detection species richness, Physical. Richness.Adult.Only= species richness represented only by adult anurans with physical captures, Physical.Richness.with.Metamorphs= species richness of all anurans detected with physical methods, Total.Richness= species richness detected by all methods

Community comparisons—There were no significant correlations of auditory and physical captures over time found between dissimilarity matrices using Mantel tests. The highest correlation was found between presence-absence dissimilarity matrices of capture and auditory methods (Mantel's $r=0.155$, $P=0.071$). Mean weekly density of physical

methods and auditory presences were less correlated (Mantel's $r=0.094$, $P=0.177$). The final comparison of error-adjusted mean weekly calling presences and mean captures was the least correlated of the three comparisons (Mantel's $r=0.038$, $P=0.346$).

Species-level comparisons—The distributions of species presences during weeks with calling showed some associations. *Bufo americanus* was found more often than expected in the breeding season than outside of the breeding season in 2011 but not significantly more than the expected distribution ($\chi^2_r=5.219$, $df=1$, $P=0.054$). The difference in distribution of capture presences of *B. fowleri* during weeks when calling was significantly more than the expected ($\chi^2_r=6.109$, $df=1$, $P=0.024$). *Bufo fowleri* was captured during all seven weeks (100%) when calling and four of the 14 weeks (28.6%) when not calling. Including captures from all methods, there was no significant difference in the distribution of capture presences during weeks with calling than the expected for *Pseudacris crucifer* ($\chi^2_r=2.524$, $df=1$, $P=0.173$). However, when only PVC pipe captures were evaluated (one transect capture during calling season), all *P. crucifer* captures were not within weeks of calling, differing significantly from the expected distribution ($\chi^2_r=5.25$, $df=1$, $P=0.048$).

Further evaluation of potential correlations within breeding seasons was investigated. *Bufo fowleri* calls at the UGRBP were detected in the range of May 11–July 8 2011, allowing for a comparison of 35 days and nine weeks. Within their calling season, *B. fowleri* captures were significantly correlated with calling activity on a weekly scale (Spearman's $r=0.799$, $n=9$, $P=0.010$) but not a daily scale (Spearman's $r=0.307$, $n=35$, $P=0.073$). On a daily and weekly scale, *B. fowleri* was generally found on days and weeks when they were calling but were never captured on a date when *B. fowleri* was

not calling (Figure 1.5 and 1.6). There was no correlation in ranks for *Hyla chrysoscelis* auditory and visual detections on a daily (Spearman's $r=-0.25$, $n=45$, $P=0.872$) or weekly scale (Spearman's $r=-0.11$, $n=17$, $P=0.966$). *Hyla chrysoscelis* showed no trend as sometimes captures increased with calling and other times decreased (Figure 1.7). No evident trend between calling and physical captures was seen in other species.

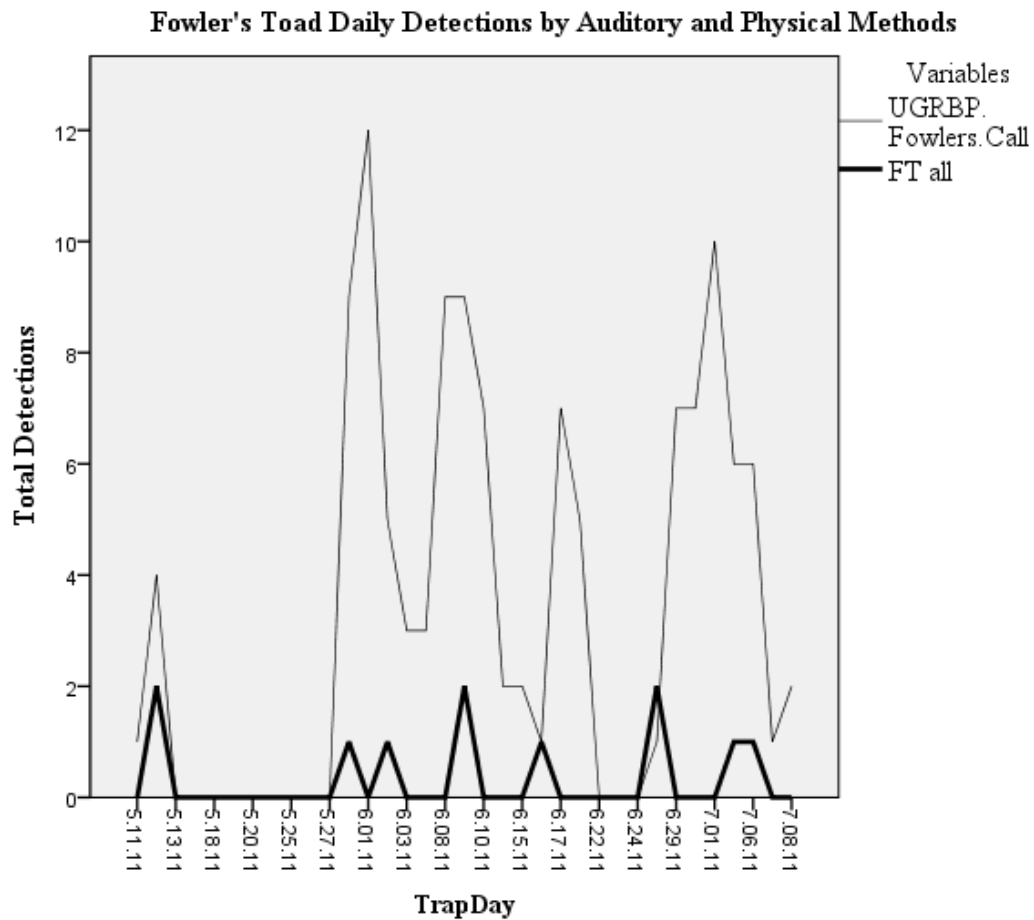


FIGURE 1.5 Comparison of the total daily detections of *Bufo fowleri* (Fowler's toad) during its breeding season using auditory and physical methods. UGRBP Fowlers.Call= the number of recordings with a detection of a Fowler's toad call at the Upper Green River Biological Preserve (UGRBP); FT all= the number of Fowler's toad captures at the UGRBP

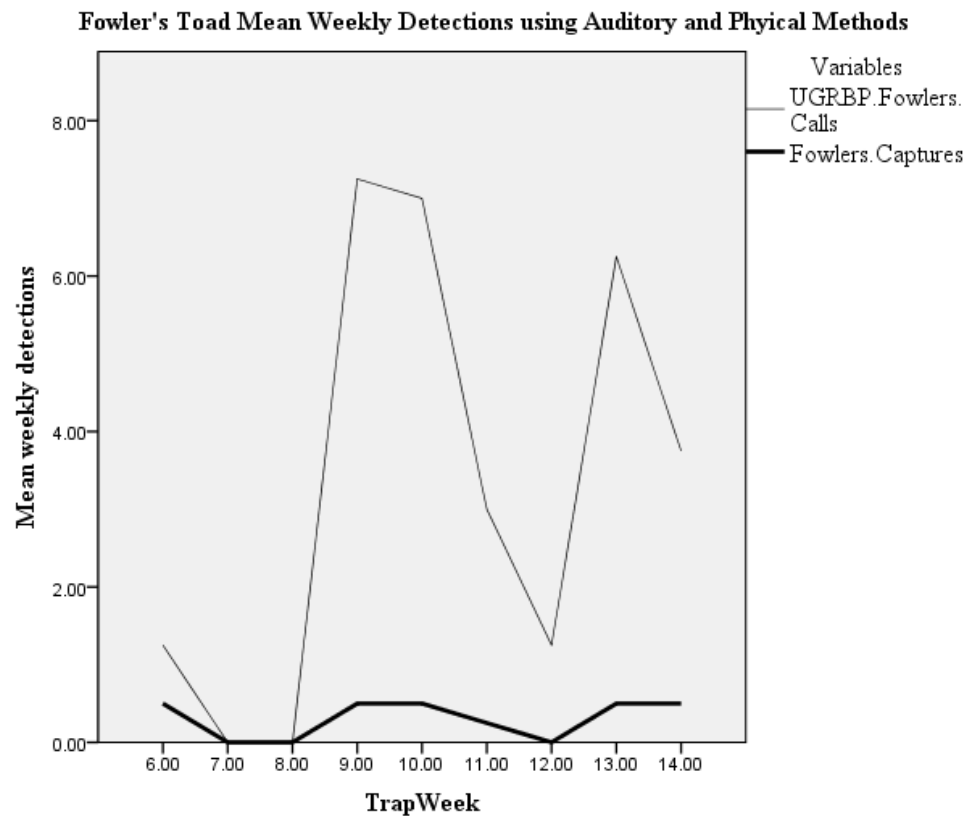


FIGURE 1.6 Comparison of the mean weekly detections of *Bufo fowleri* (Fowler's toad) during its breeding season using auditory and physical methods. UGRBP Fowlers.Call= the mean number of recordings with a detection of a Fowler's toad call at the Upper Green River Biological Preserve (UGRBP); FT all= the mean number of Fowler's toad captures at the UGRBP per week when calling

Number of Detections of *H. chrysoscelis* (Cope's Gray Treefrog) during trap weeks

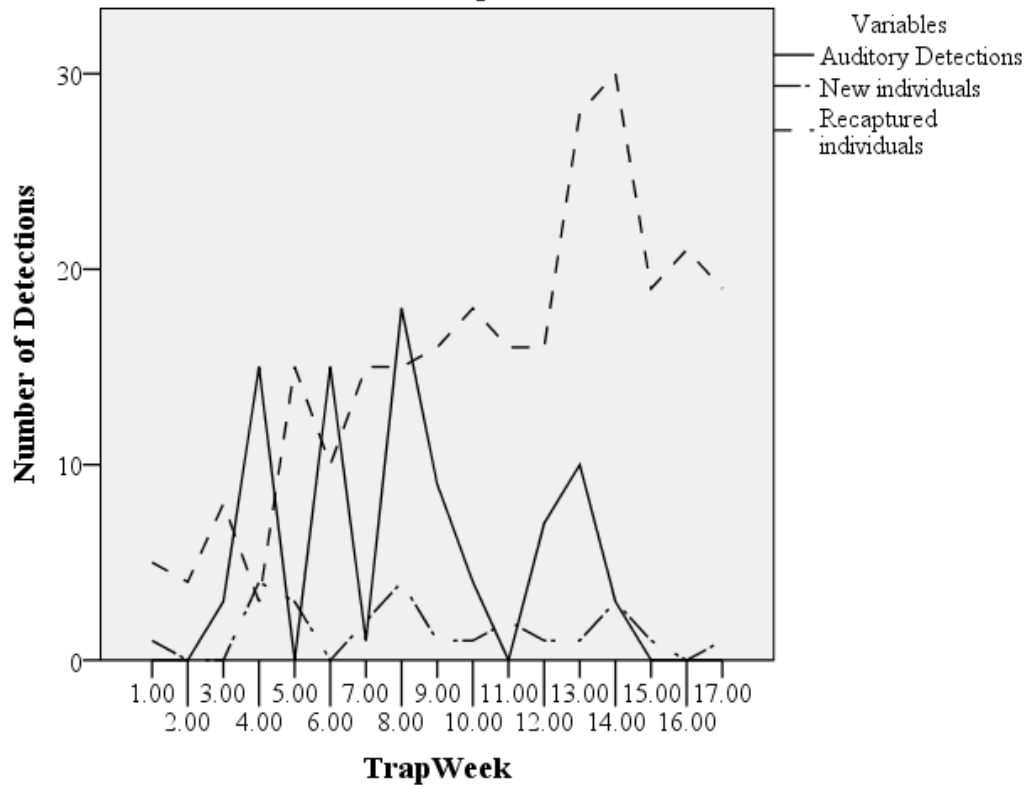


FIGURE 1.7 The total number of detections of *Hyla chrysoscelis* (Cope's gray treefrog) as observed by first capture of individuals, recaptured individuals, and the number of recordings (3 min.) with call presences. Auditory detections= the total number of recordings with *H. chrysoscelis* present during the trap week (n=4 days), New individuals= total *H. chrysoscelis* individuals not previously captured; Recaptured individuals=total *H. chrysoscelis* individuals that had previously been captured

Time of day comparisons—A comparison of the proportion of calling activity during periods of the day revealed a general correlation between WPP and the UGRBP, but the proportion of calls at the WPP were less during daytime periods. A significant correlation in ranks was found between calling activity at WPP and UGRBP for all the species evaluated including *Bufo americanus* (Spearman's $r=0.941$, $n=6$, $P=0.005$), *Hyla*

chrysoscelis (Spearman's $r = 0.897$, $n=6$, $P=0.015$), and *P. crucifer* (Spearman's $r = 0.971$, $n=6$, $P=0.001$). Visual comparison of the proportion of calling activity during each period showed a similarity in that for each species, although the proportion of calling activity was always lower (often absent) at WPP during periods the daytime periods two, three, and four (0600–1700 hrs) (Figure 1.8, 1.9, and 1.10).

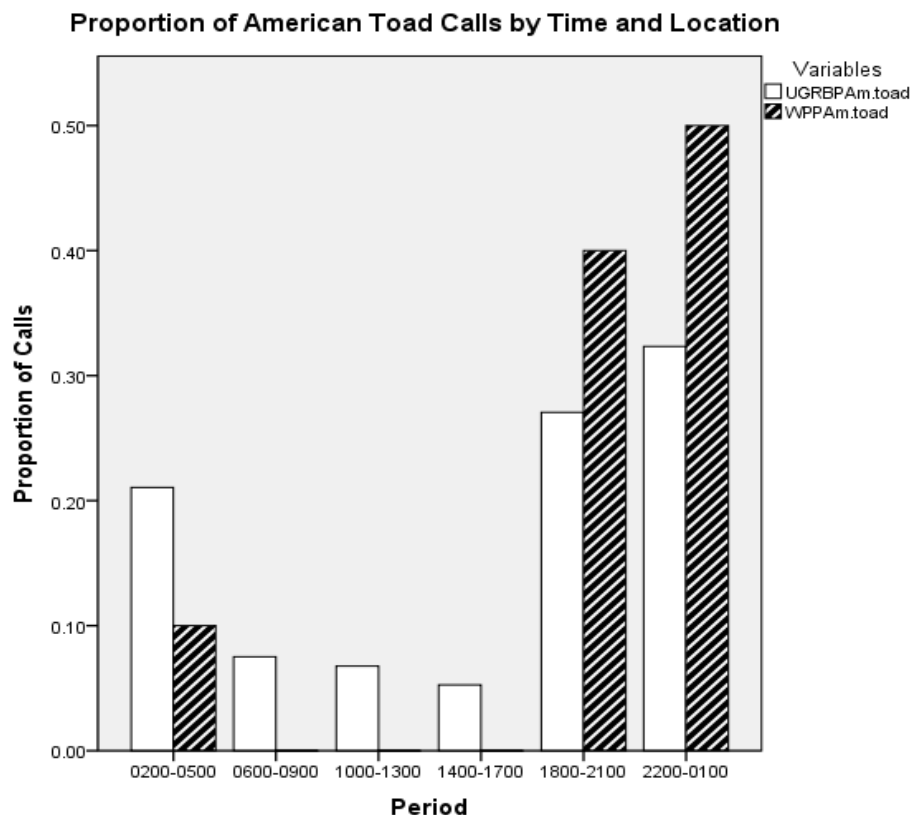


FIGURE 1.8 Comparison of the proportion of calling presence during each period of the day of *Bufo americanus* (American Toad) at the Upper Green River Biological Preserve (UGRBP) and Weldon Peete Park (WPP). UGRBP Am.toad= proportion of calls of the American toad at the UGRBP; WPP Am.toad=proportion of calls of the American toad at WPP

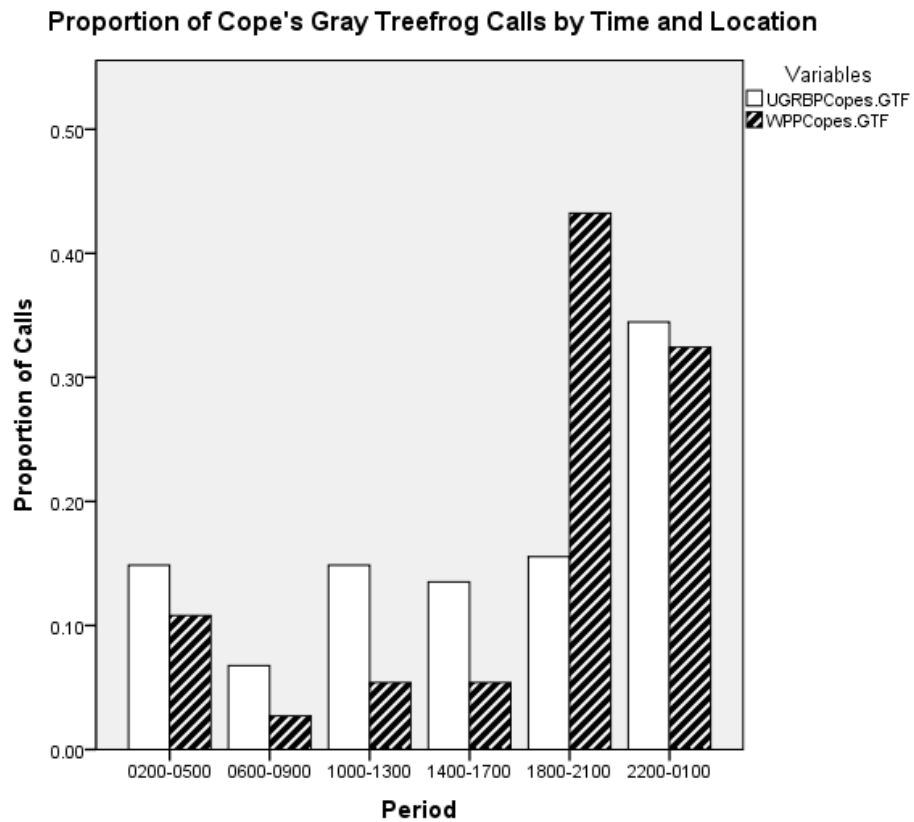


FIGURE 1.9 Comparison of the proportion of calling presence during each period of the day of *Hyla chrysoscelis* (Cope's Gray Treefrog) at the Upper Green River Biological Preserve (UGRBP) and Weldon Peete Park (WPP). UGRBPCopesGTF= proportion of calls of the Cope's gray treefrog at the UGRBP; WPPCopesGTF=proportion of calls of the Cope's gray treefrog at WPP

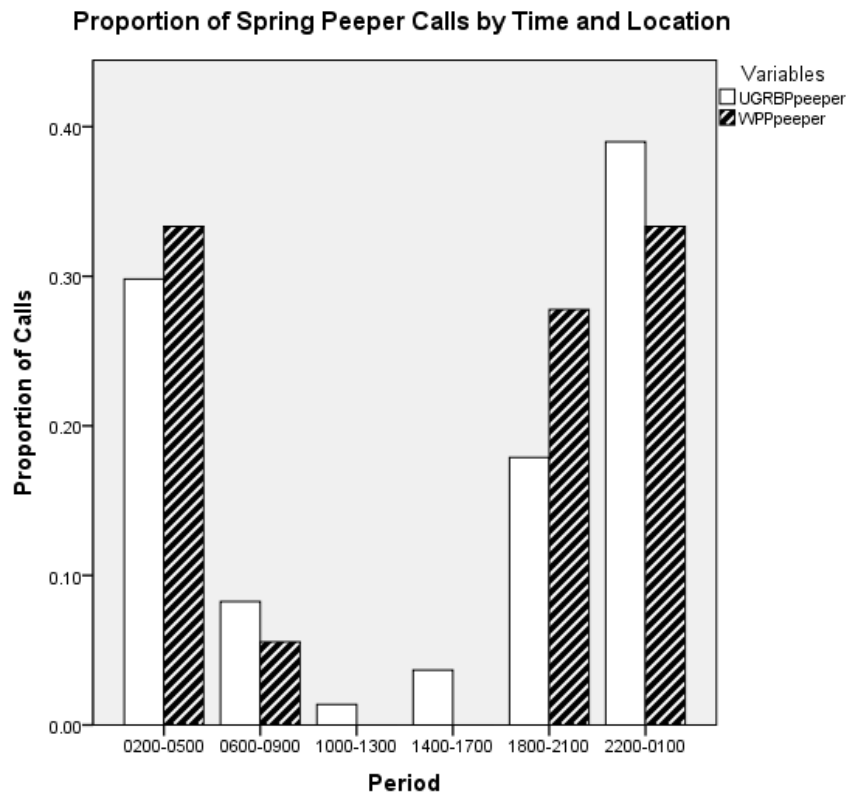


FIGURE 1.10 Comparison of the proportion of calling presence during each period of the day of *Pseudacris crucifer* (spring peeper) at the Upper Green River Biological Preserve (UGRBP) and Weldon Peete Park (WPP). UGRBPpeeper= proportion of calls of the spring peeper at the UGRBP; WPPpeeper=proportion of calls of the spring peeper at WPP

Calling season comparisons—Calling activity on dates before 23 June 2011 varied by location within WPP and the UGRBP for the three species evaluated. *Bufo americanus* was detected at the UGRBP from 6 April–11 May 2011 whereas calls were only detected from 6 April–13 April 2011 at WPP (Figure 1.11). Dates of calling for *P. crucifer* at the UGRBP were primarily from 5 April–5 May 2011 with a second period of decreased calling activity from 24 May–27 May 2011. *Pseudacris crucifer* was noted to call at WPP primarily from 8 April–15 April 2011 with a second instance of a few calls

on 14 June 2011 and 15 June 2011 (Figure 1.12). *Hyla chrysoscelis* calls were detected at the UGRBP from 19 April–22 June 2011 (through 9 July 2011 if not limited until 23 June 2011). At WPP, calls were more sporadically detected from 9 April–9 June 2011 (Figure 1.13)

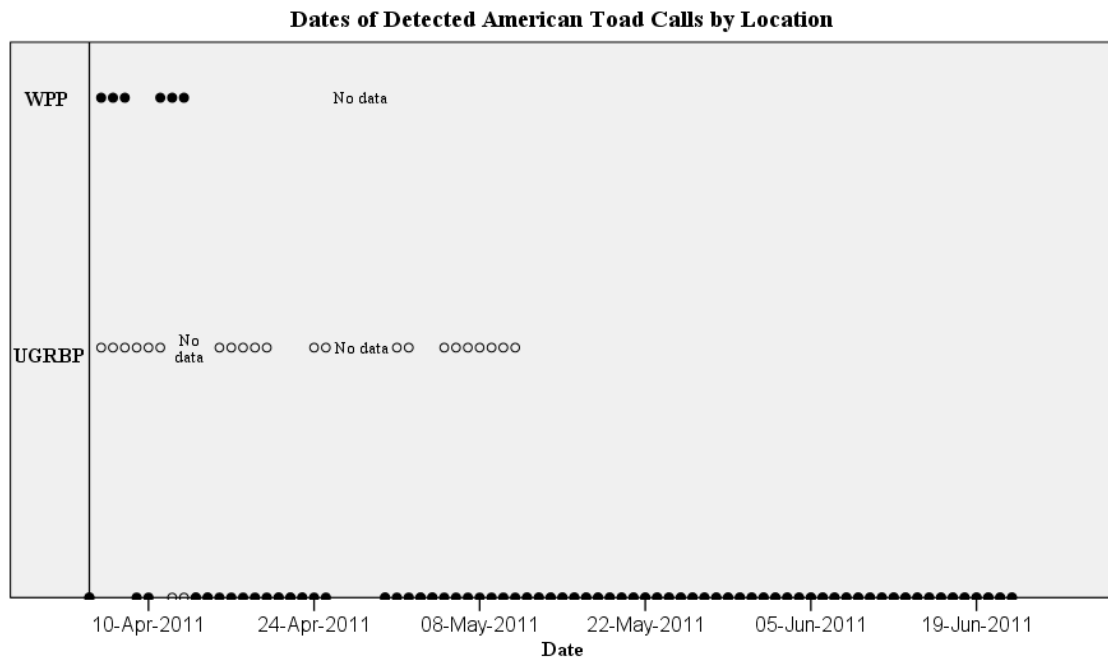


FIGURE 1.11 Dates of detected calls for *Bufo americanus* (American toad) at the Upper Green River Biological Preserve (UGRBP) and Weldon Peete Park (WPP). Open circles represent dates of calling at the UGRBP; Solid circles represent dates of calling at WPP

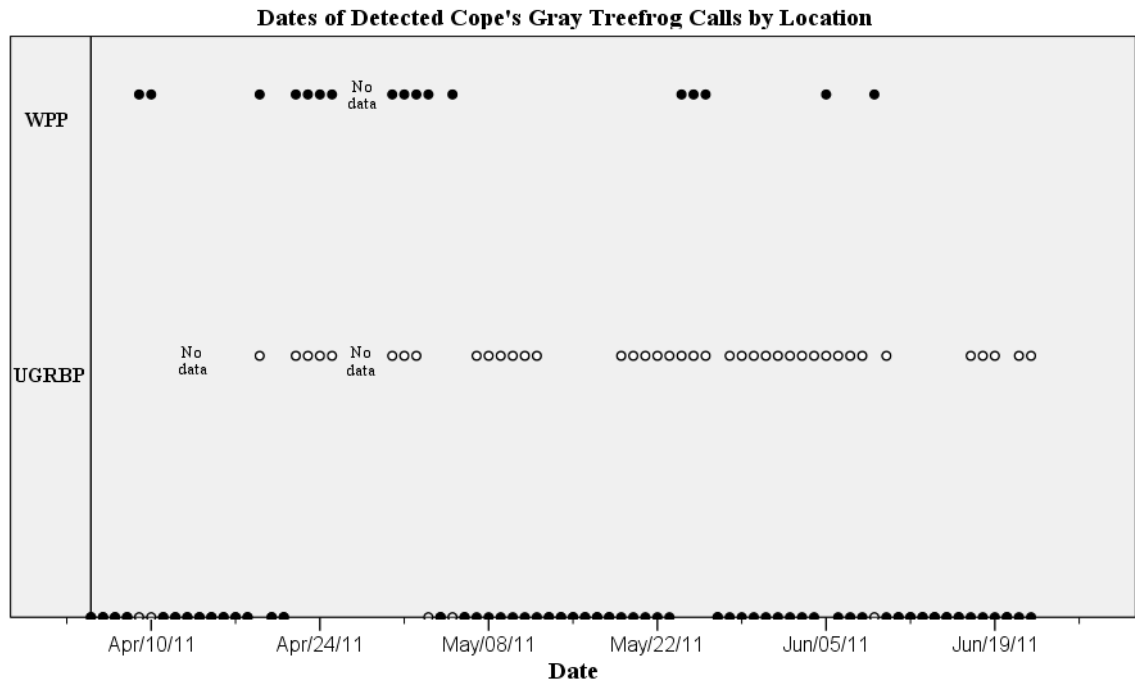


FIGURE 1.12 Dates of detected calls for *Hyla chrysoscelis* (Cope's gray treefrog) at Upper Green River Biological Preserve (UGRBP) and Weldon Peete Park (WPP). Open circles represent dates of calling at the UGRBP; Solid circles represent dates of calling at WPP

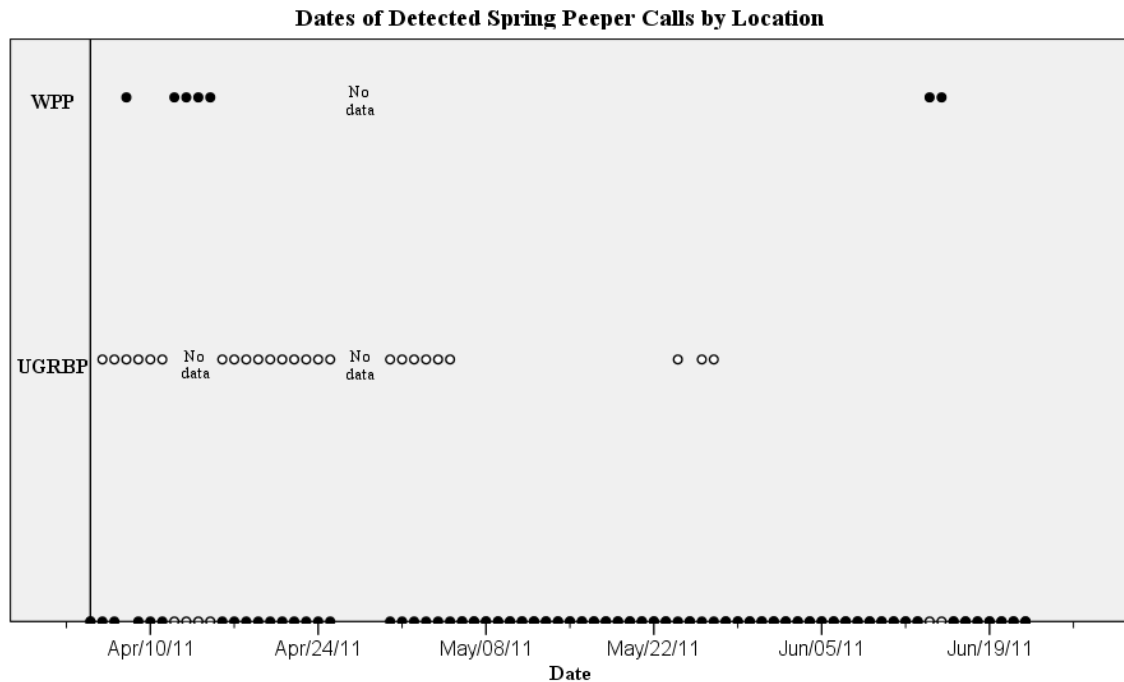


FIGURE 1.13 Dates of detected calls for *Pseudacris crucifer* (spring peeper) at Upper Green River Biological Preserve (UGRBP) and Weldon Peete Park (WPP). Open circles represent dates of calling at the UGRBP; Solid circles represent dates of calling at WPP

Influences of background amplitude—The potential effects of sound pollution were analyzed using the characteristics of vocalizations from WPP and the UGRBP. First, the vocalization length and maximum frequency was tested for correlation with the background amplitude to determine possible call plasticity. For *Bufo americanus* vocalizations at WPP, the maximum amplitude at the time of the vocalization had a significant positive relationship with the length of the call (Table 1.5). Contrastingly, the call length at the UGRBP had a significant negative relationship with background amplitude for *B. americanus* (Table 1.5). There was no relationship between amplitude and maximum calling frequency for *B. americanus* at WPP or the UGRBP (Table 1.5).

TABLE 1.5 A summary of three species of anuran's call characteristics in response to maximum amplitude of background noise. Significant relationships in linear regressions are marked with a (*). Call length was measured in seconds (s) and frequency was measured in hertz (Hz). Weldon Peete Park (WPP) is a park polluted by urban noises and the Upper Green River Biological Preserve (UGRBP) is a large biological preserve

Species	Location	Characteristic	Relationship	Statistic
<i>Bufo americanus</i>	WPP	Call length	Positive $R^2=0.401$	$F_{(1,13)}=8.72$, $P=0.011^*$
		Frequency	None $R^2=0.082$	$F_{(1,13)}=1.160$, $P=0.301$
	UGRBP	Call length	Negative $R^2=0.368$	$F_{(1,13)}=7.58$, $P=0.016^*$
		Frequency	None $R^2=0.22$	$F_{(1,13)}=3.67$, $P=0.078$
<i>Hyla chrysoscelis</i>	WPP	Call length	None $R^2=0.054$	$F_{(1,13)}=0.742$, $P=0.405$
		Frequency	None $R^2=0.012$	$F_{(1,13)}=0.154$, $P=0.701$
	UGRBP	Call length	Positive $R^2=0.339$	$F_{(1,13)}=6.654$, $P=0.023^*$
		Frequency	Negative $R^2=0.448$	$F_{(1,13)}=10.549$, $P=0.006^*$
<i>Pseudacris crucifer</i>	WPP	Call length	Negative $R^2=0.448$	$F_{(1,13)}=10.546$, $P=0.006^*$
		Frequency	None $R^2=0.001$	$F_{(1,13)}=0.005$, $P=0.943$
	UGRBP	Call length	None $R^2=0.012$	$F_{(1,13)}=0.159$, $P=0.697$
		Frequency	None $R^2=0.004$	$F_{(1,13)}=0.053$, $P=0.822$

There was no significant relationship between *Hyla chrysoscelis* call length and amplitude at WPP (Table 1). However, a significant positive relationship was found between maximum amplitude and *H. chrysoscelis* call length at the UGRBP (Table 1.5). No significant relationship between the call maximum frequency and maximum amplitude was observed at WPP (Table 1.5). A significant negative relationship was

found between *H. chrysoscelis* maximum call frequency and maximum amplitude during the time of the call at the UGRBP (Table 1.5).

A significant negative relationship was found between the length of *Pseudacris crucifer* vocalizations and the maximum amplitude during the period of the call at WPP (Table 1.5). No significant relationship was found between call length or frequency of *P. crucifer* with maximum amplitude at the UGRBP (Table 1). The relationship between call frequency and maximum amplitude was not significant at WPP (Table 1.5).

Comparisons of call characteristics by location—The differences in sound characteristics at urban (WPP) and rural (UGRGP) sites were tested for each species. For *Bufo americanus*, there was a significant effect of location on both maximum amplitude and RMS amplitude during the time of the call (Table 1.6). There was no significant effect of location on the maximum frequency and length of call for *B. americanus* (Table 1.6).

TABLE 1.6 A summary of the linear models for each species testing the difference in call characteristics and background amplitude between an urban noise-polluted location, Weldon Peete Park (WPP) and the natural noises of the Upper Green River Biological Preserve (UGRBP). RMS Amp.= Root mean square amplitude, Call Length in seconds, Max. Freq.=Maximum Frequency of call in Hertz

Species		RMS Amplitude	Max. Amplitude	Call Length (s)	Max. Freq. (Hz)
<i>Bufo Americanus</i>	UGRBP-Mean	975.5	4925.8	9.11	1856.4
	WPP-Mean	1815.1	8486.1	8.51	1912.5
	Difference	F=15.265, df=1, P=0.001*	F=6.406, df=1, P=0.017*	F=0.364, df=1, P=0.551	F=2.829, df=1, P=0.104
<i>Hyla chrysoscelis</i>	UGRBP-Mean	864.5	3897.5	1.156	2903.9
	WPP-Mean	2143.1	6053.7	0.679	2725.6
	Difference	F=34.003, df=1, P<0.001*	F=6.884, df=1, P=0.014*	F=13.806, df=1, P=0.001*	F=2.420, df=1, P=0.131
<i>Pseudacris crucifer</i>	UGRBP-Mean	895.7	3221.5	0.19	3118.5
	WPP-Mean	1548.4	4193.2	0.16	2859.0
	Difference	F=7.179, df=1, P=0.012*	F=1.591, df=1, P=0.218	F=5.051, df=1, P=0.033*	F=26.576, df=1, P<0.001*

Calls of *H. chrysoscelis* were significantly longer at the UGRBP than at WPP (Table 1.6) (Figure 1.14). Call frequency was not significantly different at WPP than at the UGRBP (Table 1.6). Significantly higher RMS amplitude but not maximum amplitude was found at WPP than the UGRBP during the evaluated calls of *H. chrysoscelis* (Table 1.6).

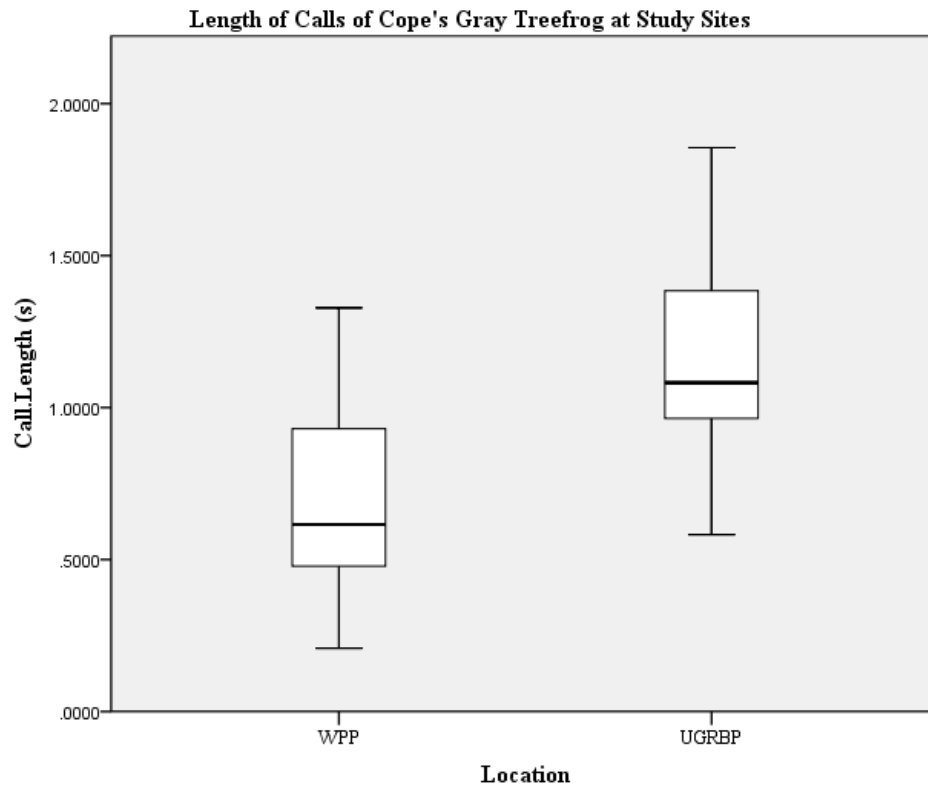


FIGURE 1.14 Comparison of the call length of *Hyla chrysoscelis* (Cope’s gray treefrog) at Weldon Peete Park (WPP) and the Upper Green River Biological Preserve (UGRBP). The upper and lower whiskers represent the range, the box represents the 25–75 percentiles, and the line within the box is the median

Pseudacris crucifer calls were significantly longer at the UGRBP than at WPP (Table 1.6) (Figure 1.15). The maximum frequency of vocalizations was significantly higher at the UGRBP than at WPP (Table 1.6) (Figure 1.16). RMS amplitude at the time of *P. crucifer* calls was significantly higher at WPP, and maximum amplitude was higher at WPP but not significant (Table 1.6).

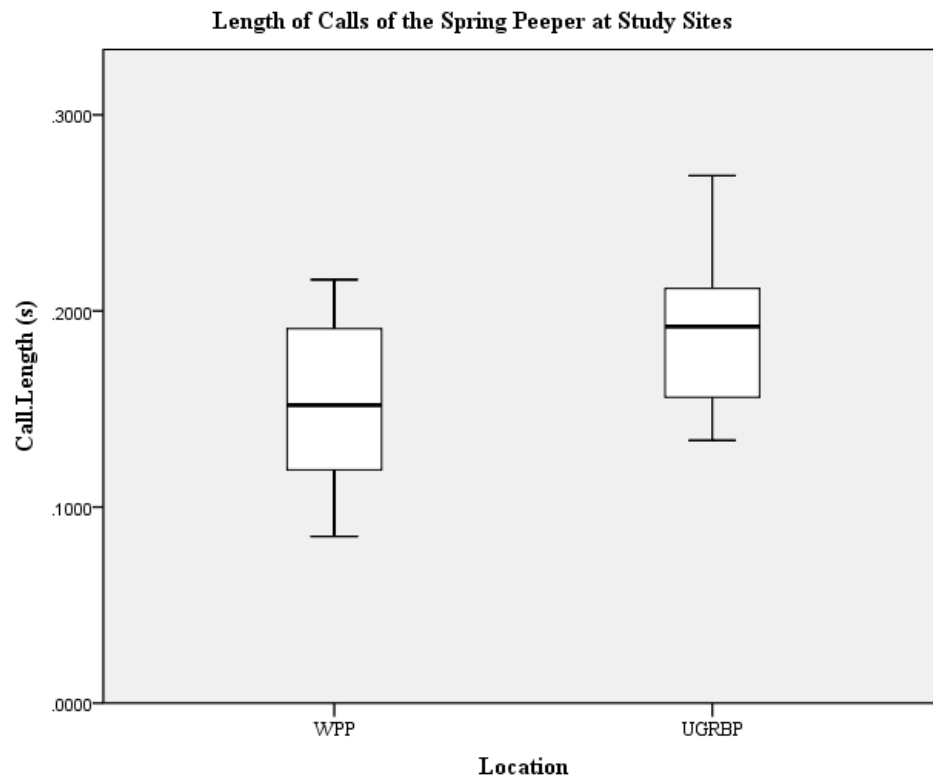


FIGURE 1.15 Comparison of the call length *Pseudacris crucifer* (spring peeper) at Weldon Peete Park (WPP) and the Upper Green River Biological Preserve (UGRBP). The upper and lower whiskers represent the range, the box represents the 25–75 percentiles, and the line within the box is the median

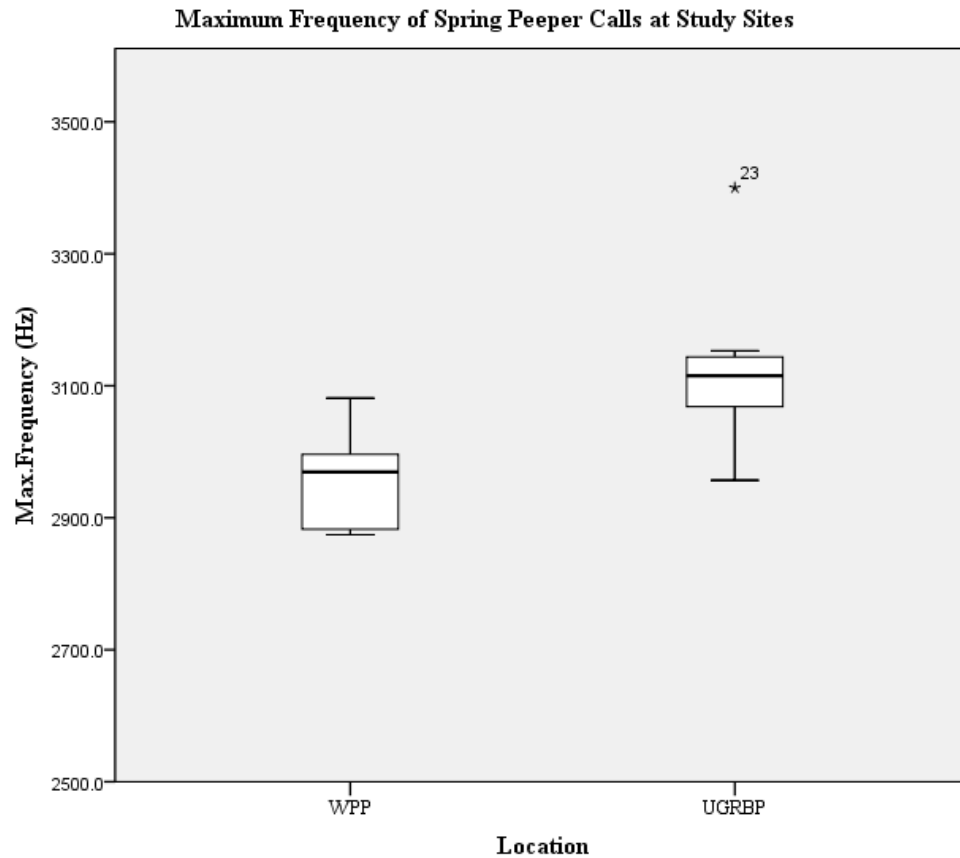


FIGURE 1.16 Comparison of the maximum call frequency (Hz) of *Pseudacris crucifer* (spring peeper) at Weldon Peete Park (WPP) and the Upper Green River Biological Preserve (UGRBP). The upper and lower whiskers represent the range, the box represents the 25–75 percentiles, and the line within the box is the median, the star is an outlier

DISCUSSION

Method comparisons— The hypothesis that auditory and physical methods were correlated was not supported by this study. The resulting communities detected by auditory and physical methods were independent of each other. However, correlations were detected on a species level. Therefore, a tradeoff exists in choice of anuran

monitoring methods depending upon the targeted species or the type of information desired by the monitoring effort.

Physical methods comparison—An overall comparison of the three physical methods (drift fences, PVC pipe array, and transects) suggested that drift fences were the most effective method to capture anuran individuals, followed by PVC pipe arrays, and visual encounter transect surveys. The success of drift fences compared to other physical methods mirror the results of other studies. Drift fences with pitfall traps were shown to be more effective for anuran captures than time-constrained searches (Crosswhite et al. 1999). Likewise, drift fences can capture more anurans and represent greater species richness than cover boards or time-constrained searches (Ryan et al. 2002). Drift fence captures in this study were possibly reduced due to the flooding of pitfall traps after heavy rains. Holes were not drilled into the bottom of pitfall buckets due the regular high water table and flooding of fields. This problem was also noted in a similar study with water-saturated environment (Dawson and Hostetler 2008). Over one-third of traps were not viable on approximately 11 of the 102 trap nights due to flooding. Furthermore, it appeared that pitfall traps were tampered by unknown animals on 19 days in June and July of 2011. Raccoons, two that were caught in baited live mammal traps along drift fences, may have been responsible for this problem. Additionally, two minks that were observed in the trap area during the summer of 2011 were possible intruders. There were cases when anurans were found in traps that appeared to be tampered, thus, depending on the invading organism, there may not have been a negative effect on anuran captures. These factors were assumed to be part of normal study error and not play a major role in evaluation but should be noted for their potential effect.

As expected, PVC pipe arrays documented little species richness (n=2) due to their restriction to arboreal anurans. However, the PVC pipe array provided the most captures (n=44) and recaptures (n=290) of a single species, *Hyla chrysoscelis*. The mean number of captures for an *H. chrysoscelis* individual was 7.6 captures and the highest number of captures for an individual was 38 total captures. Although this study employed fewer pipes, results from PVC arrays were similar to that of Boughton and Staiger (2000). The duration of time from the original capture of an individual to its last recapture ranged from one day to approximately one year. Unlike other studies wherein *P. crucifer* was absent (Boughton and Staiger 2000), individuals were present to a lesser extent (n=6) in PVC pipes as were additional metamorphs of this species. *Pseudacris crucifer* was not captured in PVC pipes in Boughton and Staiger's (2000) study or the previous studies that they referenced, but it was captured in later studies using PVC pipe methods different from those incorporated in this study (Johnson and Semlitsch 2003; Liner, Smith and Golladay 2008). Anurans in PVC pipes showed high site fidelity as only two individuals were found in different pipes (the closest bordering pipe approximately 25m away) than the original capture location. This trend in high site fidelity was also seen by Boughton and Staiger (2000) and by Pitmann et al. (2008). In the Pitmann et al. (2008) study, only three of the 82 captured *H. chrysoscelis* were captured in PVC pipes other than the original location.

Species richness comparison—In general, physical survey methods produced higher species richness than auditory methods during the course of this study. Auditory detections were limited in this survey because of a limited group of recognizer models (n=8) were constructed with AVRS. A larger repertoire of recognizer files or manually

listening to recordings would likely result in detecting higher species richness as seen in other studies (Acevedo and Villanueva-Rivera 2006). A total of 12 species were seen throughout the physical survey, and only 8 species were detected using auditory methods when recordings were available. During periods when both methods were used, 10 adult anuran species were captured using physical methods (11 including metamorphs). Seven species were detected using auditory methods during the same dates. It is likely that other species were calling during this time but were not heard in the preliminary search for vocalizations. Some species, such as *L. palustris*, are known for vocalizations that are difficult to detect (Lepage, Courois and Daigle 1997). *L. sylvatica* may have also bred before recording began. Additionally, my initial scan of recordings may not have been thorough enough to detect all of the species that were calling.

Community comparisons—When species with recognizer models were compared in a distribution throughout time (Mantel test), auditory and physical methods were independent of one another. Therefore, a study employing only either physical methods or auditory methods would detect a different anuran community composition of their study site. Part of the variation between methods is likely due to the multiple physical capture methods used in this study. Transect and pitfall traps probably better reflected the activity of some species, but PVC pipe captures were necessary to provide adequate captures of *H. chrysoscelis* and *P. crucifer* in this study and would be useful for tree frog captures in other studies. Additionally, calling activity best reflects breeding activity and not necessarily the presence of individuals for non-reproductive purposes. This is problematic as correlations between monitoring programs using either method would not

be applicable to each other. A species by species comparison brings more light to possible reasons for the lack of correlation in methods observed in this study.

Species-level comparisons; Lithobates catesbeiana and Lithobates clamitans—

The only species with a recognizer model developed that was detected only by physical methods during trap days was *L. clamitans*. Only one *L. clamitans* was recorded on 1 May 2011, a date when traps were not open. Additionally, the two individuals captured in 2011 were smaller than what is considered sexually mature and were likely not calling. A similar reason might explain the low abundance of *L. catesbeiana* at the study site. The two individuals of this species captured in July 2010 were not measured, but they were likely smaller than what is considered sexually mature for *L. catesbeiana*, thus were not calling. Both *L. clamitans* and *L. catesbeiana* have breeding habitats in permanent slow-moving, small streams, or still water (Harding 1997) all of which were not represented at the study area, therefore the small number of calls and the absence of adults is not surprising. Few calls and captures were detected for *L. catesbeiana* and *L. clamitans*, thus both methods acceptably predicted a low population of these species.

Acris crepitans—The abundance of *A. crepitans* vocalizations was also less than what was expected from physical captures. The only two recordings with vocalizations detected of this species were on 25 May 2011, the day after the last individuals were captured using physical methods. The physical presence of *A. crepitans* was noted regularly throughout April and May of 2011. Breeding choruses are known to begin in early April in Southern Illinois (Gray 1983), therefore it is unclear why individuals were not calling at this location. It is unclear if their breeding habitat was within the temporary pools formed during the early spring or on the banks of the river as both would be within

the expected habitat for this species (Harding 1997). Using only physical methods, it is likely that one would assume this area was a breeding location for *A. crepitans*. Nevertheless, the lack of calling activity and the absence of metamorphs during both years suggest that the study site was not used primarily as a breeding location for this species. Only physical methods were adequate to detect the extent of the *A. crepitans* population, consequently, they would be preferable to monitor *A. crepitans* at this location.

Lithobates palustris and *Scaphiopus holbrookii*—Species that were not detected with auditory methods but were captured using physical methods included *L. palustris*, *L. sphenoccephala*, and *S. holbrookii*. It should be noted that even with 19L (5 gallon) bucket pitfall traps, I observed that large adult individuals of *L. palustris* and *L. sphenoccephala* were able to escape, thus lowering captures of these species. It does not appear that *L. palustris* used the study area as a breeding site because there were no vocalizations noted in recordings or metamorphs observed in physical surveys, but individuals are also known to breed earlier in the year before recordings began (Conant and Collins 1998). It is also not apparent if *S. holbrookii* was breeding at this location during the study period as no metamorphs or vocal activity were noted. *Scaphiopus holbrookii* is known to be an explosive breeder in temporary or permanent pools during heavy rains such as those found at this location (Punzo 1992). However, the only two captures of this species were recorded on 29 June 2010 and 29 July 2010. Both dates were during times when there was no standing water in the fields, making it unlikely that this movement was for reproduction.

Lithobates sphenoccephala—*Lithobates sphenoccephala* was known to breed at this location because an abundance of individuals of this species were observed in various stages of metamorphosis in the temporary pool. The tadpoles of *L. sphenoccephala* take 50–75 days to complete metamorphosis (Wright 1932). Tadpoles were largely developed into metamorphs by June 7 2011 when the pools dried, thus the primary breeding period was likely in March, prior to the period assessed in this study. Unless effective recognizer models were developed, physical survey methods are more effective for *Lithobates sphenoccephala*.

Pseudacris feriarum and *Lithobates sylvatica*—The only species not observed by physical methods that were detected in the recordings was *P. feriarum*. *Pseudacris feriarum* is known to have an early breeding season lasting from late December to April depending upon the geographic region (Conant and Collins 1998), therefore it is likely that more vocalizations would have been heard if recordings from dates before April were analyzed. The probability of capture by physical methods would therefore also increase on earlier dates. Metamorphs but no adults of *L. sylvatica* were found, and this could have been a consequence of the timing of the survey as this species breeds earlier in the year (Conant and Collins 1998). Metamorphs of *P. feriarum* were observed at the study site even though breeding appeared to have occurred at this location. The cause of this discrepancy is unknown, but failure of physical methods to detect metamorphs and adults is possible. Therefore, AVRS should be used to monitor *P. feriarum* individuals at this location.

Bufo americanus—The *B. americanus* captures were not significantly higher during weeks of calling, but three of the four weeks with capture presences were during

weeks with detected vocalizations. It appears likely that the four individuals captured during the breeding season were a result of migration to the temporary pool for reproduction, and the single individual may have been a more permanent resident of the area. Migration to a breeding site has previously been described for *B. americanus* for distances up to approximately 600 m (Oldham 1966). Regardless, the overall lack of individuals was surprising given the number of calls recorded. This showed that at this type of location, auditory methods might be more effective for monitoring the use of a breeding location.

Pseudacris crucifer—Captures of *P. crucifer* were few and were generally outside of the calling season. As with *Hyla chrysoscelis*, the majority of captures was with PVC pipes and might not reflect movement. Supporting this hypothesis, the sole capture of an adult individual in a week with calling was an individual caught in a transect by the temporary pool in field A. Otherwise, the absence of *P. crucifer* in the PVC pipes suggested that their use was not beneficial for reproduction. Both *P. crucifer* and *H. chrysoscelis* individuals may have used the PVC pipes to decrease desiccation in high temperatures and low humidity environments. The ability of PVC refuges to reduce desiccation of inhabitants has previously been suggested (Johnson 2005). Therefore, it is more likely that the use of the PVC pipe array for *P. crucifer* was for physiological reasons and does not reflect breeding activity. The use of auditory methods to monitor populations for this species is probably preferable over physical sampling unless a more effective method for the capture *P. crucifer* is developed.

Bufo fowleri—The distribution of capture presences of *B. fowleri* was related to the weeks when calling activity was detected. Not surprisingly, a correlation of physical

and auditory detections was also observed for *B. fowleri* during the breeding season. More individuals were captured during trap weeks when calling activity was present. This correlation suggested that conditions that were not suitable for calling during the breeding season were also not ideal for movement. Similar overall patterns in movement in relation to calling activity were seen in a *B. fowleri* population in Connecticut as more adult individuals were captured during periods of greater calling intensity (Clark 1974). There appeared to be no distinct advantage to either method during the breeding season at this location because *B. fowleri* was easily captured and monitored using AVRS. Nevertheless, a greater abundance of data was obtained at this location using AVRS, providing a more informative view for long-term management.

Hyla chrysoscelis—Movement of *H. chrysoscelis* appeared to represent a more complicated trend as a correlation in captures and calling was not significant. Peaks in captures sometimes positively trended and other times negatively trended with spikes in calling activity (Figure 1.21). Part of this inconsistency is likely due to the capture method (PVC pipe array) that serves as both long-term and short-term refuges. One notable trend occurred during the most intense week of calling in 2011 (May 24–May 27). During this week, recaptures declined as new individuals increased, suggesting a large movement of individuals at the study site. A similar trend was seen in a PVC array study of *H. chrysoscelis* where recaptured individuals decreased and new individuals increased during the beginning of the breeding season (Pittman et al. 2008). Longer term studies would be necessary to further clarify this potential trend. The use of auditory methods for *H. chrysoscelis* would be best used to monitor breeding activity whereas

physical methods should be effective for monitoring population trends not related to breeding activity.

Sound pollution—The exploratory comparisons of auditory data from WPP and the UGRBP supported the hypothesis that vocalizations would change in response to background noise amplitude and that vocalization characteristics would be different between locations. An evaluation of auditory data from WPP and the UGRBP noted sizable differences in species composition and prevalence. Of the eight species with an AVRS recognizer model available, only four species were found at WPP compared to the eight at the UGRBP. Species with no vocalizations detected at WPP included *Acris crepitans*, *Bufo fowleri*, *Lithobates catesbeiana*, and *Pseudacris feriarum*. Additionally, fewer recordings with anurans vocalizations present were detected for all species at WPP (n=67) than the UGRBP (n=893). *Pseudacris feriarum* has been known to adapt well to agricultural areas and is attracted to flooded fields (Jensen et al. 2008) but was not detected at WPP. The temporary pools at WPP may have dried too soon for *P. crucifer* to successfully breed in this location. The lack of a still permanent body of water suitable for reproduction may explain the absence of *L. catesbeiana* and only a single detection of *L. clamitans* at WPP as noted at the UGRBP. The absence of *A. crepitans* calls at WPP is likely due to the lack of a potential breeding location at WPP. Whereas WPP likely provides adequate habitat requirements, *A. crepitans* has also been noted for population declines due to drought, fertilizer use, highway salts, and pollution (Harding 1997) that are likely more prevalent at WPP. Perhaps the most unexpected call absence was that of *B. fowleri*. *Bufo fowleri* commonly calls near river banks (Conant and Collins 1998). Even though the full breeding season was not represented in recordings (only until 23

June 2011), *B. fowleri* began calling on 11 May 2011 at the UGRBP, thus recording dates were probably not significant factors. Many similarities exist between the Barren River at WPP and the Green River at the UGRBP, but the Barren River at this location could have been less suitable for breeding for *B. fowleri* due to less abundant shallow areas and pooling in the river.

The relative lack of abundance of anurans at WPP could be due to other factors. First, an adequate breeding location may not have been available in the vicinity of either recorder. Pooling water was present in both fields and wooded areas in the spring, but data were not collected to determine the length of time these pools existed before drying. Because anuran calling is largely for reproductive purposes (Gerhardt and Huber 2002), it would be expected that less calling activity would be present at locations not used for breeding. Second, the effects of the bordering urban environment were having a significant negative impact on anuran populations at WPP. Even when permanent breeding ponds are available, anuran species richness and abundance may be reduced in urban landscapes compared to agricultural or forested landscapes (Gange and Fahrig 2007). The cause of decreased populations may be due to a variety of negative impacts related to urban landscapes such as habitat fragmentation, introduction of invasive species, and increased pollutants (Bradley 1995). Physical surveys would be helpful in providing a clearer picture of the status of anuran populations at this site.

Time of day comparisons—The calling activity based on periods of the day showed differences between calling locations. All species called more frequently at evening and night at both WPP and the UGRBP. However, at WPP, there was always a lesser proportion of calls during the daytime periods ranging from 0600–1700. This trend

may be due to increased noise pollution in urban environments during daytime hours, reducing calling activity, whereas less competition with noise would always be present for individuals at a rural location.

Calling season comparisons—The dates of calling seasons were also different at each location. The calling season generally appeared to be shortened at WPP compared to the UGRBP, but this may have been an artifact of the fewer calls detected at WPP. Furthermore, the absence of recordings from before April 4 2011 prohibits most speculation as to the timing of the breeding season. Interestingly, from graphical representation of calling seasons, there may be a shift of calling at WPP toward earlier in the season. A temperature-induced change seems unlikely due to close proximity of sites (<50 km), but water temperatures and air temperatures may be inflated due to its proximity to the city. The most clear possible shift in a breeding season was seen with *Hyla chrysoscelis*. Still, more data would need to be collected, especially prior to April, to determine any relationships of dates with reasonable certainty.

Influences of amplitude and location on call characteristics—The comparison of anuran vocalization characteristics at WPP and the UGRBP revealed a variety of trends for the three species evaluated, *Bufo americanus*, *Hyla chrysoscelis*, and *Pseudacris crucifer*. In general, both the maximum and RMS amplitude during the time of the vocalization was greater at WPP than the UGRBP. The source of the background noise at each site also likely affected vocalization characteristics differently. Background noise at WPP during the time of call was generally unnatural such as traffic noise and the “hum” of the water treatment plant. In contrast, the greatest source of background noise at the UGRBP was from natural sources and varied by season. When present, *P. crucifer*

choruses were a dominant nighttime sound in the early spring whereas in later spring and summer, insect sounds were most prevalent. In most cases, the primary noise sources at WPP were at lower frequencies than anuran vocalizations as is common in an urban environment (Warren et al. 2006). Background sounds at the UGRBP were at higher frequencies or within the same range of anuran calls.

The vocalizations of *Bufo americanus* showed opposing trends at each site in regard to call length. Vocalizations at the UGRBP were longer when the maximum background amplitude was lower. Conversely, call length at WPP was longer during the presence of higher background amplitude. Further evaluation of call length at both locations revealed why this disparity may have been present. The amplitude range of 5,000–10,000 was the range wherein calls were shortest in both locations. When amplitude was below this range (occurring only at the UGRBP), calls were longer. When amplitude was above this range (occurring only at WPP), calls were lengthened. Call length has been shown to be independent of size and highly variable among individuals of *B. americanus* (Howard and Young 1998), suggesting call plasticity of this call characteristic. Due to the relatively quieter vocalization of *B. americanus*, quiet conditions at the UGRBP (in general, quieter *Pseudacris crucifer* choruses) might have allowed longer vocalizations to effectively communicate information whereas in noisier conditions shorter calls could efficiently communicate in gaps of competing anuran vocalizations. Conversely, vocalizations at WPP were nearly always competing with unnatural, higher amplitude background noise (10,000–25,000) and therefore generally calls were shorter. During exceptionally high amplitude noise interference, call length may have been increased to improve the odds of calls being heard even in exceptionally

high noise. Call duration has been noted to change to a moderate extent when presented with high amplitude background noises, consistent with this study (Penna et al. 2005). Although not statistically significant, mean call maximum frequency was higher at WPP than the UGRBP, also indicating a possible difficulty in communication in an environment with low frequency noise. Higher call frequencies in the presence of urban or traffic noise have been noted for a variety of species including *Bufo woodhouseii* (Woodhouse's toad), *Hyla cinerea* (green treefrog), *Litoria wingii* (brown treefrog), and *Crinia signifera* (common eastern froglet) (Parris et al. 2009; Barrass 1986).

The vocalizations of *Hyla chrysoscelis* were found to be variable at the UGRBP but not at WPP. The observed trend was that as the maximum background amplitude increased, the call length increased and maximum call frequency decreased. Possibly, these trends are evident at the UGRBP due to breeding competition among individuals. Calls that are longer and lower may indicate fitness and maturity to a potential mate in a competitive environment (Arak 1983). A lower frequency could also have been needed to distinguish calls from *Pseudacris crucifer* choruses that occupied the frequency range just above that of *H. chrysoscelis*. This hypothesis is also supported by the frequency of vocalizations being significantly lower at the UGRBP than WPP. There would be less incentive for lower frequencies at WPP because there was lessened competition with other males and no *P. crucifer* chorus. Additionally, heightened frequencies may be a result from competition with consistently higher amplitude background noise (Parris et al. 2009).

The call length of *Pseudacris crucifer* was affected by the maximum amplitude of background noise at WPP but not at the UGRBP. Calls length at WPP decreased when

maximum amplitude increased. It is unclear why this trend is present as one might predict that calls would be longer in response to noisier environment. Alternatively, calls may be shorter in order to be more distinct from contrasting background noises.

Vocalizations at the UGRBP were at higher frequencies and longer than those at WPP. This trend is likely due to the highly competitive anuran choruses present at the UGRBP but not at WPP. The frequency range of *P. crucifer* is usually higher than that of other species at the UGRBP, excluding those of *P. feriarum*. Therefore, lower frequency calls may be masked by other calls and higher frequency calls are more likely to not be in competition with other species. A longer call in chorus activity is also necessary as there is an increased chance of shorter calls being masked by other species. Sexual selection studies of *P. crucifer* show female preference for lower frequency calls (Forester and Czarnowski 1985), thus lessening the advantage for high frequency male calls.

Recommendations—Because there is not a clear correlation between physical and auditory methods, less potential exists to effectively monitor anurans on a large scale. This study shows clear advantages to using either auditory or physical survey methods. Often, the most appropriate method may be dependent upon the species of interest in the study (Table 1.7). Other times, a specific type of data is desired for a study that may dictate survey methods (Table 1.8). Categories and recommendations for specific purposes are based on those evaluated for other breeding habitats (Muths and Iko 2000). Advantages of auditory recording have improved since these earlier recommendations resulting from a methods comparison (Muths and Iko 2000) including digital recording capability, less frequent maintenance, better reliability of recording devices, and automated detections of recordings (decreasing time spent listening). Observations such

as those made in this study that employ both methods may also be preferable for particular studies or monitoring programs. Future studies should attempt to factor calling density into auditory methods and construct recognizer models for all species potentially found at the study location. Evaluating habitats with a greater quantity of anurans such as ponds should provide a larger data source and allow for more comparisons.

TABLE 1.7 Recommendations of monitoring methods for each species observed in this study at the site of this study and other study sites (based on this study and others cited).

Auditory methods assumes the use of AVRS taking development of recognizers into account; when both methods acceptable (1)=preferred based on this studies efforts

Species	UGRBP Study Site	Other study sites
<i>Acris crepitans</i>	Physical	Physical (1) or Auditory
<i>Bufo americanus</i>	Auditory	Physical or Auditory (1)
<i>Bufo fowleri</i>	Physical or Auditory (1)	Physical or Auditory (1)
<i>Gastrophyrne carolinensis</i>	Not enough information	Not enough information
<i>Hyla chrysoscelis</i>	Physical (1) or Auditory	Physical (1) or Auditory
<i>Lithobates catesbeiana</i>	Physical	Physical or Auditory (1)
<i>Lithobates clamitans</i>	Physical	Physical (1) or Auditory
<i>Lithobates palustris</i>	Physical	Physical (1) or Auditory
<i>Lithobates sphencephala</i>	Physical	Physical (1) or Auditory
<i>Pseudacris crucifer</i>	Auditory	Auditory
<i>Pseudacris feriarum</i>	Auditory	Auditory
<i>Lithobates sylvatica</i>	Not enough information	Not enough information
<i>Scaphiopus holbrookii</i>	Physical	Not enough information

TABLE 1.8 Recommendations of methods based on study design or purpose. Categories partially adapted from Muths and Iko (2000) and columns copied are noted with (*).

ARS= automated recording system, MCS=manual calling survey, Intensive= intensive search survey, AVRS= Automatic Vocalization Recognition software, Physical= physical methods in this study REC=recommended, ACC=acceptable, INA=inappropriate, EXP=expensive, AFF=affordable

Data or study design	ARS*	MCS*	Intensive*	AVRS	Physical
Species with audible calls	REC	ACC	ACC	REC	ACC
Species with weak/no calls ¹	INA	INA	REC	INA	REC
Estimating abundance	POOR	POOR	REC	POOR	REC
Relative abundance	POOR	ACC ²	ACC	ACC ²	ACC
Behavior and phenology	REC	POOR	POOR	REC	POOR
Reliability	ACC	ACC	REC	REC	ACC
Short-term cost ²	EXP	AFF	EXP	EXP	ACC
Long-term cost ²	NA	NA	NA	ACC	EXP
Single-species ecology	REC	POOR	REC	REC	REC
multiple species; small area ³	REC	POOR	REC	REC	REC
Multiple species; large area ³	POOR	ACC	ACC	ACC	ACC
¹ Concerning the monitoring of all amphibian species					
² Acceptable for some species					
³ Geographic area					

The evaluation of the potential effects of sound pollution highlights the potential uses of ADRS and AVRS for questions outside of those normally considered in amphibian conservation. Anurans in this study appeared to show call plasticity to a degree that was also observed by Cunningham and Fahrig (2010). This study could be improved if a greater number of species were evaluated on a greater time scale (years). A future study would preferably include species that both do and do not have call plasticity. It is hypothesized that populations that cannot change call characteristics may decline over time (Sun and Narins 2005). From this study, it appears that calls of all anuran species studied were affected by noise pollution. Noise pollution should be a concern in

anuran conservation as a change in calling behavior reflects the attempt of individuals to change their natural behavior to a modified environment, potentially reducing the fitness of an individual or species.

APPENDIX 1

Physical captures of other herpetofauna and small mammals—Other herpetofauna noted during the physical methods in the study included 94 captures representing 13 species (Table 1.9). *Eumeces* sp. was not identified to species level due to the slight differences that differentiate the three possible species within the area. The only individual identified to the species level from this genus was *Eumeces fasciatus* (five-lined skink). The majority of species were captured either in drift fence pitfall traps or on the drift fence and included *Ambystoma jeffersonianum* (Jefferson salamander), *Ambystoma maculatum* (spotted salamander), *Chelydra serpentina* (snapping turtle), *Eumeces* sp. (five-lined, SE five-lined, or broadhead skink), *Eurycea lucifuga* (cave salamander), *Notophthalmus meridionalis meridionalis* (red-spotted newt), *Storeria dekayi dekayi* (Northern brown snake), and *Storeria occipito maculata* (redbelly snake). *Eumeces* sp. was often observed basking on the black fabric of the drift fence. Snakes captured in pitfall traps were smaller than those observed in transects. Herpetofauna only observed in transects were *Graptemys geographica* (common map turtle), *Opheodrys aestivus* (rough green snake), *Terrapene carolina carolina* (Eastern box turtle), and *Thamnophis sirtalis sirtalis* (Eastern garter snake).

Small mammals totaling 213 captures were not identified to species level but simply noted as a mouse, vole, or shrew and were categorized by their family names Cricetidae (mice and voles) and Soricidae (shrews). All captures of small mammals in

this study were accomplished through drift fence pitfall traps. A single juvenile *Sylvilagus floridanus* (Eastern cottontail) was also captured in a pitfall trap. Additionally, mammals caught in live traps from 06/23/2011–07/29/2011 included two *Procyon lotor* (common raccoon), one *Didelphis virginiana* (Virginia opossum), and one *Marmota monax* (groundhog).

TABLE 1.9 A list of all herpetofaunal and mammal captures during physical surveys at the Upper Green River Biological Preserve. Animals were not identified as individuals, therefore the number of captures may represent recaptures.

Herpetofauna species	Number of captures
<i>Ambystoma jeffersonianum</i> ¹ (Jefferson salamander)	3
<i>Ambystoma maculatum</i> (spotted salamander)	1
<i>Chelydra serpentina</i> (snapping turtle)	1
<i>Eumeces</i> sp. ² . (five-lined, SE five-lined, or broadhead skink)	41
<i>Eurycea lucifuga</i> (cave salamander)	8
<i>Graptemys geographica</i> (common map turtle)	2
<i>Notophthalmus meridionalis meridionalis</i> ³ (red-spotted newt)	4
<i>Opheodrys aestivus</i> (rough green snake)	1
<i>Scincella lateralis</i> (ground skink)	4
<i>Storeria dekayidekayi</i> (Northern brown snake)	12
<i>Storeria occipitomaculata</i> (redbelly snake)	5
<i>Terrapene carolinacarolina</i> (Eastern box turtle)	11
<i>Thamnophis sirtalis sirtalis</i> (Eastern garter snake)	1
Mammal species	
Family Cricetidae ⁴ (mice and voles)	49
Family Soricidae ⁴ (shrews)	164
<i>Sylvilagus floridanus</i> (Eastern cottontail)	1
<i>Didelphis virginiana</i> ⁵ (Virginia opossum)	1
<i>Marmotamonax</i> ⁵ (groundhog)	1
<i>Procyonlotor</i> ⁵ (raccoon)	2
¹ Indiviudals were juveniles and thus difficult to identify with certainty ² Subtle differences distinguish species, only species identification was one <i>Eumeces fasciatus</i> ³ Only observed in red eft stage ⁴ Small mammal species were generally not identified to species level ⁵ Mammals captured during 6/23/11–7/29/11 as part of effort to remove animals tampering with drift fences	

Summary of dates and times of recordings based on recorder location:

TABLE 1.10 Summary of recording dates and times used for analysis by the automatic vocalization recognition software, Song Scope[®]. UGRBP= Upper Green River Biological Preserve, WPP= Weldon Peete Park

Trap Period	UGRBP field A (BP01,BP04) ¹	UGRBP field B (BP02,BP03) ¹	WPP Site 1	WPP Site 2
June-July 2010	BP01 7/15: 1100– 7/29: 0900 (n=359)	BP02 7/15: 1100– 7/27:1000 (n=282)	NA	NA
October 18– 29, 2010	BP01 10/18: 1000– 10/29: 0900 (n=264)	BP01 10/18: 1000– 10/29: 0900 (n=264)	NA	NA
April-July 2011	BP04 4/04: 0900– 4/12:1200 (n=194) BP04 4/15: 1300 – 4/26: 1000 (n=262) BP04 4/29: 0900– 6/27: 1400 (n=1422) BP01 6/27: 1500– 7/04: 0800 (n=162) BP01 7/04: 0900– 7/29: 0900 (n=601)	BP03 4/04: 0900– 4/12:1200 (n=194) BP03 4/15: 1000 – 4/15: 1200 (n=3) BP02 4/15: 1300– 4/26: 1000 (n=264) BP03 4/29: 0800– 6/27: 1300 (n=1422) BP02 6/27: 1400– 7/04: 0800 (n=163) BP03 7/04: 0900– 7/29: 0900 (n=601)	BG01 4/04: 0900 – 4/26: 1200 (n=530) BG01 4/29: 0900– 6/23: 1400 (n=1326)	BG02 4/04: 0900 – 4/26: 1200 (n=530) BG02 4/29: 0900– 6/23: 1400 (n=1326)
Total Recordings	3264	3193	1856	1856
¹ BP01 and BP04 are replicate recorders at the same location; BP02 and BP03 are replicate recorders at the same location				

Distribution of captures at the UGRBP—The UGRBP study site was divided into four trapping areas. Each area included one drift fence, eight PVC pipes, and half of the transects (n=2) located in that field. Area 1 included the half of the field A furthest from the river, and Area 2 included the half closest to the river. Areas in field B were

designated as Area 3, including the half of the field furthest from the river, and Area 4, including the half closest to the river. The distribution of the first capture of adult anurans among the trapping areas were assessed using a chi-square test. The study site was also assessed for anuran distribution by testing the distribution of adult anurans in field A to field B and side 1 (Area 1+ Area 3) and side 2 (Area 2+ Area 4) using Chi-square tests. The effectiveness of the three physical trapping techniques (drift fences, transects, and PVC pipe arrays) was compared based on their ability to capture new anuran individuals using a Chi-square test.

The distribution of original adult captures in the four field areas was not significantly different than the expected distribution ($\chi^2=3.612$, $df=3$, $n=147$, $P=0.306$). Additionally, the distribution of original captures was not significant considered by each field ($\chi^2=0.007$, $df=1$, $n=147$, $P=0.934$) or by each side ($\chi^2=2.456$, $df=1$, $n=147$, $P=0.117$). When the distribution of original captures based drift fence captures only were compared, significant differences were found from the expected distribution in both the area of the fence ($\chi^2=11.89$, $df=3$, $n=72$, $P=0.008$) and side of the field where the fence was located ($\chi^2=9.389$, $df=1$, $n=72$, $P=0.002$). The distribution of anurans based on the fence location was not significantly different from the expected distribution ($\chi^2=2.0$, $df=1$, $n=72$, $P=0.157$). The distribution of original captures by the PVC pipe array according to the area was found to be significantly different than the expected distribution ($\chi^2=10.837$, $df=3$, $n=49$, $P=0.013$). PVC pipe capture distribution in each field (field A=34, field B=15) was also significantly different than the expected distribution ($\chi^2=7.367$, $df=1$, $n=49$, $P=0.007$), but the distribution according to each field side was not significantly different from the expected distribution ($\chi^2=2.469$, $df=1$, $n=49$,

$P=0.116$). The distribution of original captures was not significantly different than expected in transect areas ($\chi^2=3.0$, $df=3$, $n=25$, $P=0.392$), fields ($\chi^2=1.96$, $df=1$, $n=25$, $P=0.162$), or sides ($\chi^2=1.0$, $df=1$, $n=49$, $P=0.317$).

The distributions of original adults captured were relatively even among each area of the UGRBP, but the distributions of original captures by each physical method (drift fence, PVC pipe array, and transects) were more informative. Reflecting the overall trend in capture rates, there were no differences in the distribution of anurans captured in transects. Not surprisingly, captures varied by its location at the study site as the majority of drift fence captures occurred on side 1 representing fences closest to the river (side 1=49, side 2=23). The only species with successful reproduction (metamorphs captured) directly observed in the temporary pool in field A were *L. sphenoccephala* and *P. crucifer*. Other species observed within or around this pool during respective breeding seasons included *A. crepitans*, *B. americanus*, *L. catesbeiana* (a single, dead individual), and *P. feriarum* (heard calling near but not seen at pool). Even though these species likely bred in this pool, it is possible that other breeding sites were more important for some terrestrial species, explaining why there was no difference in individuals captured in drift fences between fields. For instance, pools with plentiful metamorphs (>200) of *B. fowleri* were observed on an island in the Green River below field B. The river almost certainly also served as breeding locations or home ranges for other terrestrial anuran species. The use of river riparian zones as both breeding and non-breeding habitat has been noted for a variety of species (Panik and Barrett 1994; Burbrink et al. 1998). The higher concentration of anurans captured in drift fences near the river is likely due to this reason. Contrary to my expectations, there were not a greater number of anurans

captured in field A even though it had more potential breeding locations and cover than field B (A=30, B=42). Apparently, the field with greater amount of habitat heterogeneity and cover did not influence the preference of anurans for this location. Perhaps this disparity is because drift fences were located in edge habitat. Anurans could have avoided being in the comparatively open field B and thus preferentially used the edge habitat for movement.

The distribution of PVC pipe adult individual captures was concentrated in field A (A=34, B=15). *Hyla chrysoscelis* individuals (n=44) accounted for the majority of the PVC pipe captures, and *P. crucifer* (n=6) was the only other species caught using this method. Activity was particularly prevalent in pipe #7 and #8 where a total of 13 individuals (30%) were found. Two *H. chrysoscelis* individuals in pipe #8 were observed in amplexus. The locations of these two pipes were the closest to the temporary pool (<20m) on the side furthest from the river. Additionally, *H. chrysoscelis* vocalizations were more frequently heard in field A when traps were being checked. A close relative to *H. chrysoscelis*, *Hyla versicolor* (gray tree frog) has been shown to stay in close proximity (<60m) to its breeding site during the breeding season (Johnson and Semlitsch 2003). The preference of *H. chrysoscelis* to stay in a close proximity to a breeding site in a PVC pipe array was also noted by Pitmann et al. (2008). The temporary pool in field A is a likely breeding location for *H. chrysoscelis* and an observed breeding location for *P. crucifer*, explaining the preference of arboreal species for this field. Due to the use of habitat nearby pools, this study re-emphasizes the need for protection of lands nearby breeding locations of *Hyla* populations.

BIBLIOGRAPHY

- Acevedo, M.A., and L.J. Villanueva-Rivera. 2006. Using Automated Digital Recording Systems as an Effective Tool for the Monitoring of Birds and Amphibians. *Wildlife Society Bulletin* 34:211–214.
- Arak, A. 1983. Male-male competition and mate choice in anuran amphibians. Pp. 181–205 *In* *Mate Choice*. Gordon, P. (Ed.). Cambridge University Press, Oxford, USA.
- Barrass, A.N. 1986. The effect of highway noise on phonotactic and associated reproductive behaviour of selected anurans. Ph.D. Dissertation, Vanderbilt University Nashville, Tennessee, USA.
- Boughton, R.G. and J. Staiger. 2000. The use of PVC pipe refugia as a sampling technique for Hylid treefrogs. *American Midland Naturalist* 144:168–177.
- Bowker, R.G., and M.E. Cambron. 2006. An automated digital recording system: the amphibulator. *Multimedia* 8:592–600.
- Bradley, G.A. (Ed.) 1995. *Urban forest landscapes: incorporating multidisciplinary perspectives*. University of Washington Press, Seattle, USA.
- Burbrink, F.T., C.A. Phillips, and E.J. Heske. 1998. A riparian zone in southern Illinois as a potential dispersal corridor for amphibians and reptiles. *Biological Conservation* 86:107–115.
- Clark, R.D. 1974. Activity and Movement Patterns in a Population of Fowler's Toad, *Bufo woodhousei fowleri*. *American Midland Naturalist* 92:257–274.
- Conant, R., & Collins, J.T. 1998. *A Field Guide to the Reptiles and Amphibians of Eastern and Central North America*. Third Edition, Expanded. Houghton Mifflin Company. New York, New York, USA.
- Corn, P.S., E. Muths, and Iko, W.M. 2000. A Comparison in Colorado of Three Methods to Monitor Breeding Amphibians. *Northwestern Naturalist* 81:22–30.
- Cross, C. L. 2000. A new design for a lightweight squeeze box for snake field studies. *Herpetological Review* 31:34.
- Crosswhite, D.L., S.F. Fox, and R.E. Thill. 1999. Comparison of Methods for Monitoring Reptiles and Amphibians in Upland Forests of the Oachita Mountains. *Oklahoma Academy of Science* 79:45–50.

- Crump, M.L., and Scott, N.J. 1994. Standard techniques for inventory and monitoring: visual encounter survey. Pp. 85–92 *In* Heyer W. R., M. A. Donnelly, R. W. McDiarmid, L. C. Hayek, and M. S. Foster. Measuring and monitoring biological diversity: Standard methods for amphibians. Smithsonian Institution Press, Washington D. C., USA.
- Cunnington, M.G. and L. Fahrig. 2010. Plasticity in the vocalizations of anurans in response to traffic noise. *Acta Oecologia* 36:363–370.
- Dawson, D.E. and M.E. Hostetler. 2008. Herpetofaunal use of edge and interior habitats in urban forest remnants. *Urban Habitats* 5:103–125.
- Dodd Jr., K. 2003. Monitoring Amphibians in Great Smoky Mountains National Park. Washington D.C., USGS Survey Circular.
- Dorcas, M.E., and C.R. Peterson. 1994. Automated Data Acquisition. Pp. 47–57 *In* Measuring and Monitoring Biological Diversity. Smithsonian Institution Press Washington D.C., USA.
- Dorcas, M.E., Price, S.J., Walls, S.C., and Barichivich, W.J. 2009. Auditory Monitoring of Amphibian Populations. Pp. 281–298 *In* Amphibian Ecology and Conservation: A Handbook of Techniques. Dodd Jr., K. Oxford University Press, New York, New York, USA.
- Dunson, W.A., R.L. Wyman, & E.S. Corbett. 1992. A Symposium on Amphibian Declines and Habitat Acidification. *Journal of Herpetology* 26:349–352.
- Eigenbrod, F., S.J. Hecnar, and F. Lenore. 2009. Quantifying the Road-Effect Zone: Threshold Effect of Motorway on Anuran Populations in Ontario, Canada. *Ecology and Society*: 14:24.
- Enge, K. M. 2001. The pitfalls of pitfall traps. *Journal of Herpetology* 35:467–478.
- Forester, D.C., R. Czarnowski. 1985. Sexual selection in the spring peeper, *Hyla crucifer* (amphibia, anura): role of the advertisement call. *Behaviour* 92:112–128.
- Gange, S.A., and L. Fahrig. 2007. Effect of landscape context on anuran communities in breeding ponds in the National Capitol Region, Canada. *Landscape Ecology* 22:205–215.
- Genet, K.S. and L.G. Sargent. 2003. Evaluation of methods and data quality from a volunteer-based amphibian call survey. *Wildlife Society Bulletin* 31:703–714.

- Gerhardt, H.C. and F. Huber. 2002. Acoustic communication in insects and anurans: common problems and diverse solutions. University of Chicago Press, Chicago, Illinois, USA.
- Gray, R.H. 1983. Seasonal, annual, geographic variation in color morph frequencies of the cricket frog, *Acris crepitans*, in Illinois. *Copeia* 2:300–311.
- Greenburg, C.H., D.G. Neary, and L.D. Harris. 1994. A comparison of herpetofaunal sampling effectiveness of pitfall, single-ended, and double-ended funnel traps with drift fences. *Journal of Herpetology* 28:319–324.
- Halliday, T. R. 1996. Amphibians. In W. J. Sutherland, *Ecological Census Techniques* (pp. 205–216). Cambridge: Cambridge University Press.
- Harding, J. H. 1997. Amphibians and Reptiles of the Great Lakes Region. The University of Michigan Press. Ann Arbor, Michigan, USA.
- Hopkins, W.A. 2007. Amphibians as Models for Studying Environmental Change. *Institute for Laboratory Animal Research Journal* 48:270–277.
- Howard R.D., and J.R. Young. 1998. Individual variation in male vocal traits and female mating preferences of *Bufo americanus*. *Animal Behaviour* 55:1165–1179.
- Jensen, J.B., Camp, C.D., and W. Gibbons. 2008. Amphibians and Reptiles of Georgia. University of Georgia Press, Athens, Georgia, USA.
- Johnson, J. R. 2005. Multi-scale investigations of the gray treefrog (*Hyla versicolor*) movements: patterns and migration, dispersal, and gene flow. Ph.D Dissertation, University of Missouri, Columbia, Missouri, USA 73p.
- Johnson, J. R., & Semlitsch, R. D. 2003. Defining the core habitat of local populations of the gray tree frog (*Hyla versicolor*) based on choice of oviposition site. *Oecologia* 137:205–210.
- Legendre, P. and L. Legendre. 1998. Numerical Ecology: 2nd English Edition. Elsevier Science B.V, Amsterdam, Holland.
- Lepage, M., R. Courtois, and C. Daigle. 1997. Surveys of calling anurans in Quebec using volunteers. Pp. 128–140 *In* Amphibians in Decline: Canadian Studies of a Growing Problem, Herpetological Conservation Volume 1, Green, D. M. Society for the Study of Amphibians and Reptiles, St. Louis, Missouri, USA.
- Liner, A.E., L.L. Smith, and S.W. Golladay. 2008. Amphibian distributions in three types of isolated wetlands in southern Georgia. *American Midland Naturalist* 160:69–81.

- Lotz, A. and C.R. Allen. 2007. Observer bias in anuran call surveys. *Journal of Wildlife Management* 71:675–679.
- McCarthy, M.A. and K.M. Parris. 2004. Clarifying the effect of toe clipping on frogs with Bayesian statistics. *Journal of Applied Ecology* 41:780–784.
- Nelson, G.L. and B.M. Graves. 2006. Anuran population monitoring: comparison of the North American Amphibian Monitoring Program's calling index with mark-recapture estimates of *Rana clamitans*. *Journal of Herpetology* 38:355–359.
- Oksanen, J., F.G. Blanchet, R. Kindt, P. Legendre, P.R. Minchin, R.B. O'Hara, et al. 2011. *vegan: Community Ecology Package* R package version 2.0-0.
- Oldham, R. S. 1966. Spring movements in the American toad, *Bufo americanus*. *Canadian Journal of Zoology* 44:63–100.
- Panik, H. R., and S. Barrett. 1994. Distribution of amphibians and reptiles along the Truckee River system. *Northwest Science* 3:197–204.
- Parris, K.M., M. Velick-Lord, and J.M. North. 2009. Frogs Call at a Higher Pitch in Traffic Noise. *Ecology and Society*:14:25.
- PASW Statistics. 2009, July 30. SPSS statistical software package Release 18.0.0.
- Penna, M., H. Pottstock, and N. Velasquez. 2005. Effect of natural and synthetic noise on evoked vocal responses in a frog of the temperate austral forest. *Animal Behaviour* 70:639–651.
- Philiott, A.D., L.F. Skerratt, K.R. McDonald, F.L. Lemckert, H.B. Harry, J.M. Clarke, et al. 2007. Toe Clipping as an Acceptable Method of Identifying Individual Anurans in Mark-Recapture Surveys. *Herpetological Review* 38:305–308.
- Pittman, S.E., A.L. Jendreck, S.J. Price, and M.E. Dorcas. 2008. Habitat Selection and Site Fidelity of Cope's Gray Tree Frog (*Hyla crysoscelis*) at the aquatic-terrestrial ecotone. *Journal of Herpetology* 42:378–385.
- Punzo, F. 1992. Dietary overlap and activity patterns in sympatric populations of *Scaphiopus holbrookii*. *Florida Scientist* 55:38–44.
- R Development Core Team. R: A language and environment for statistical computing v. 2.13.1. Vienna, Austria: R Foundation for Statistical Computing.
- Raven Version 1.2.1. 2003–2004. Ithaca, New York: Cornell Lab of Ornithology.
- Royle, J.A. and W.A. Link. 2006. Generalized site occupancy models allowing for false positive and false negative errors. *Ecology* 87:835–841.

- Ryan, T.J., T. Phillipi, Y.A. Leiden, M.E. Dorcas., T.B. Wigley, and J.W. Gibbons. 2002. Monitoring herpetofauna in a managed forest landscape: effects of habitat types and census techniques. *Forest Ecology and Management* 167:83–90.
- Stuart, S.N., J.S. Chanson, N.A. Cox, B.E. Young, A.S. Rodrigues, D.L. Fischman, et al. 2004. Status and Trends of Amphibian Declines and Extinctions Worldwide. *Science*:1783–1786.
- Sun, J.W. and P.M. Narins. 2005. Anthropogenic sounds differentially affect amphibian call rate. *Biological Conservation* 121:419–427.
- Todd, B.D., C.T. Winne, J.D. Wilson, and J.W. Gibbons. 2007. Getting the drift: examination of the effects of timing, trap type, and taxon on herpetofaunal drift fence surveys. *American Midland Naturalist* 158:292–305.
- Trifa, V.M., A.N. Kirschel, C.E. Taylor, and E.E. Vallejo. 2008. Automated species recognition of antbirds in a rainforest using hidden Markov models. *Journal of the Acoustical Society of America* 123:2424–2431.
- Trimboli, S. R. 2010. A comparison of the singing acitivity of the Carolina Wren (*Thryothorus Ludovicianus*) in Urban and Rural Settings. M.Sc. Thesis, Western Kentucky University, Bowling Green, KY, USA. 98 p.
- Waddle, J.H., T.F. Thigpin, and B.M. Glorioso. 2009. Efficacy of Automatic Vocalization Recognition. *Herpetological Conservation and Biology* 4:384–388.
- Warren, P.S., M. Katti, M. Ermann, and A. Brazel. 2006. Urban bioacoustics; it's not just noise. *Animal Behaviour* 71:491–502.
- Wright, A.H. (1932). *North American Salientia (Anura)* Number 2. Macmillin Press, New York, New York, USA.

**CHAPTER 2—AN ASSESSMENT OF FACTORS INFLUENCING ERROR RATES
OF BOTH HUMAN LISTENERS AND AUTOMATIC VOCALIZATION
RECOGNITION SOFTWARE IN MONITORING ANURAN POPULATIONS**

Abstract—Among declining amphibian populations, anurans are often selectively monitored because of their calling activity, but error in call identification is a serious concern in accurately tracking population trends. Manual calling surveys have been used to monitor large geographic areas by traveling to study sites. More recently, automated digital recording systems (ADRS) have allowed for an easier method to remotely monitor calling activity. Machine-learning methods, such as automated vocalization recognition software (AVRS), can be used to quickly search large auditory datasets for targeted vocalizations. Error rates of human listeners are biased, and both humans and computers may consistently produce false positive and false negative errors. I compared the error rates of the AVRS Song Scope[®] to error rates of human listener groups by using sets of recordings from multiple locations. I expected human error to vary by the tested treatments including the ADRS location, listener group, species, and season, and computer error to be uninfluenced. Human and computer results were used to cross-validate one another and assess the error rates of each method. Computer error rates were higher than human listeners, yet less affected by the evaluated treatments. All treatments influenced human error, whereas computer error was only affected by location, species, and ADRS placement. Both methods were biased, tending to produce more false positives than false negatives. This study suggests that error should be considered in any monitoring program, and the choice of a call survey method is dependent upon the objectives of the study.

INTRODUCTION

Wildlife researchers have recently documented amphibian population declines at alarming rates (Stuart et al. 2004). In large part, anthropogenic causes such as habitat modification are held responsible for these declines (Gibbs et al. 2005; Alford and Richards 1999), but many population fluctuations are the result of unknown causes (Stuart et al. 2004). Population trends are often unclear due to the absence of geographically widespread and long-term data (Blaustein et al. 1994).

Anurans (frogs and toads) are among the most vocal vertebrate groups (Gerhardt and Bee 2006), allowing for detections by distinct vocalizations. Calls range from simple tones or trills to more complex vocalizations similar to those found in higher vertebrates (Gerhardt and Huber 2002). Anuran calls are generally associated with reproduction, and in most species, only males call (Gerhardt and Huber 2002). The most common call is advertisement that achieves a variety of functions including species identity, reproductive status, status, size, and location (Wells and Schwartz 2006). Other calls include the aggressive call in among male competition (Wells and Schwartz 2006). Orientation to a sound source, phonotaxis, is expressed to a high degree in anurans (Gerhardt and Huber 2002). Call variation occurs among species as some have high amplitude calls relative to body size that may persist several hundred meters, whereas others have calls that attenuate quickly (Gerhardt and Klump 1988). Anurans' ability to communicate with vocalizations is essential for locating breeding sites and mates (Gerhardt and Bee 2006). Calling activity is essential for anurans and also allows for a more convenient way to monitor populations.

Several private and government programs have attempted to fill the need for long-term data over wide regions by implementing manual call surveys (Weir et al. 2005). Manual call surveys require participants to actively visit selected breeding locations and document the vocalizations heard during a given period. These surveys provide document the presence or absence of species as well as general population density. The largest of these programs, the North American Amphibian Monitoring Program (NAAMP), has conducted numerous anuran call surveys since 1997, using hundreds of participants in a variety of locales including a large portion of the Eastern U.S. (Weir and Mossman 2005).

Manual call surveys can be conducted with reduced cost and sampling effort compared to methods that attempt to capture anurans (Mazerolle et al. 2007). Call surveys have disadvantages because they are subject to imperfect detection by observers (McClintock et al. 2010). Previous studies have largely evaluated variation in participant error rates based on the protocol of a given manual call survey program (e.g. Lotz and Allen 2007; Pierce and Gutzwiller 2007; McClintock et al. 2010). Other research has focused on the effect of survey protocol variables (e.g. survey duration) and the detection probability on occupancy models (e.g. de Solla et al. 2005; Gooch, et al. 2006). Although manual call surveys are useful for collecting small samples of qualitative data, automated recording systems (ARS) provide an alternative method to gather more thorough calling activity information (Dorcas et al. 2009).

Since the development of microprocessor-based dataloggers in the 1970s, ARS have been used to collect data at specific times (Dorcas and Peterson 1994). ARS may provide advantages over manual calling surveys including extended sampling time (thus

a greater chance of detecting inconspicuous species), decreased disturbance of the sampling area, a permanent sampling record that may be checked by multiple sources, and a resource to examine both temporal and auditory variation in vocalizations (Bridges and Dorcas 2000). Previous studies of bird and amphibian calls suggest that digital versions of ARS, Automated Digital Recording Systems (ADRS), are more effective than traditional point and transect surveys in both quality and quantity (Acevedo and Villanueva-Rivera 2006). Furthermore, ARS allow a single researcher to review hundreds of hours of recordings in a controlled environment that is less affected by environmental factors such as wind and rain in manual call surveys (Waddle et al. 2009).

A prior study used recordings to evaluate the error rates of 179 past volunteers of the Michigan Frog and Toad Survey (MFTS) by listening to 12 five-minute tracks from breeding locations in Michigan wetlands (Genet and Sargent 2003). The volunteers' species identification significantly varied from the actual species on the tracks in 6 of the 12 recordings. Although error rates varied by species, only approximately 60% of respondents correctly identified all species (Genet and Sargent 2003). Another study focused on the type of observer error, noting that listeners in manual calling surveys preferentially committed commission errors (false positives or type I error) at two locations of 19% and 23.8% rather than omission errors (false negatives or type II error) of 1.2% and 1.8% (Lotz and Allen 2007). This is worrisome given simulation models of anuran populations predict that even small amounts of false positive and false negative errors may have considerable effects on site occupancy estimates (Royle and Link 2006). Advances in technology have attempted to improve upon the negative aspects of manual calling surveys.

The prevalence of digital recordings has led to the introduction of a variety of machine-learning methods for automatically searching recordings and identifying anuran vocalizations (e.g. Brandes et al. 2006). Only recently, automatic vocalization recognition software (AVRS) has become commercially available for personal computer use (Waddle et al. 2009). Although AVRS is relatively untested, it shows promise of reducing recording processing time and research cost along with the possibility of becoming more accurate than human listening or manual calling surveys (Dorcas et al. 2009). A study of the error rates of the AVRS Song Scope showed false positive rates of 2.7–15.8% and false negative rates of 45–51% in the identification of three species of anurans (Waddle et al. 2009). Furthermore in that study, adjusting the settings of Song Scope was shown to have a large effect on the balance of false positives and false negative errors (Waddle et al. 2009).

To examine the reliability of call survey methods, I calculated the error rates of both the AVRS program Song Scope[®] (v.3.1a) and human listeners. To do this, human listeners and ARVS evaluated the presence or absence of four anuran species from identical recordings from ADRS. I then cross-checked the data from both sources and validated the actual presence or absence of a given species. The goal of this experiment was not only to determine the bias of both methods using a single data source but also to provide an extensive validated datasets that could be used to adjust the AVRS detection rates to the desired levels. I hypothesized that the ARVS error would be homogeneous and random but the human error would be biased and therefore not homogeneous and random.

METHODS

I conducted this study in two trials using similar methods. Both studies determined error rates based on location, species, and group. The first trial (Trial I) evaluated error rates during different seasons. The second trial (Trial II) included more species and a greater variety of locations. Considering that the trials took place at different times, they will be introduced separately.

Trial I— Two ADRS, Song Meter[®] SM2 dataloggers, were deployed approximately 1.5-m high on trees on the banks of the Green River in the Upper Green River Biological Preserve (UGRBP) in Hart County, KY in 2009 (Table 2.1). The UGRBP is approximately 475 ha in size and encompasses a variety of habitats along the Green River corridor and largely consists of deciduous forest and old fields, some that are being restored to barrens. It is approximately 2 km from the Mammoth Cave National Park border and is largely surrounded by farmland outside of the preserve area. Also in 2009, two ADRS were also placed in a similar location in Weldon Peete Park within the city limits of Bowling Green, KY (Table 2.1). Bowling Green is the third largest city in Kentucky with 58,067 residents in a land area of 97.85 km² in the 2010 census (Bureau, U.S. Census. 2011. *Kentucky State and County Quick Facts*. Available from census.gov [accessed 11 November 2011]). WPP is approximately 30 ha of deciduous forests mixed with hayed fields, and a paved bike and walking path runs throughout its entirety (Trimboli 2010). The park is bordered by a main road and the Barren River, a tributary of the Green River. Nearby, there is a large hospital (0.5 km) and a railroad (0.63 km) (Trimboli 2010).

TABLE 2.1 The location, recorder name, and coordinates of recorders in Trial I. Weldon Peete Park is located in Bowling Green, KY and has the recorders BG (Bowling Green) 01 and BG 02 at this site; the Upper Green River Biological Preserve is located in Hart County, KY and BP (Biological Preserve) 01 and BP02 are located at this site

Location	Recorder	Coordinates
Weldon Peete Park	BG01	N37° 00' 0.003" W86° 25' 24.6"
	BG02	N36° 59' 58.2" W86° 25' 33.00"
Upper Green River Biological Preserve	BP01	N37° 14' 57.3" W85° 59' 29.5"
	BP02	N37° 14' 53.9" W85° 59' 13.8"



FIGURE 2.1 Aerial photograph of the recorder locations at the Upper Green River Biological Preserve. Recorders are approximately 300-m apart.



FIGURE 2.2 Aerial photograph of the recorder locations at Weldon Peete Park in the city of Bowling Green, KY. Recorders are approximately 300-m apart.

The Song Meter[®] SM2 data loggers are commercially available and consist of an automated recording platform inside waterproof housing and dual SMX-II weatherproof acoustic microphones. They recorded audio in the 20–20,000 Hz frequency range only during the 3-minutes at the beginning of every hour.

Recording dates from 2009 were selected based upon the presence of quality recordings and the ability to detect seasonal differences in calling activity. We selected and downloaded recordings from one ADRS at each location in seasons including April 13–23 (spring), June 1–11 (summer), and October 1–11 (fall). Due to the large size of the datasets, we only evaluated data from odd numbered hours on odd numbered days (e.g. June 1 = 0100, 0300,...) and even hours on even numbered days. We chose four species to study that were known to call at the study site and represented acoustically

distinct calls: *Bufo americanus* (American toad), *Bufo fowleri* (Fowler's toad), *Pseudacris crucifer* (spring peeper), and *Pseudacris feriarum* (upland chorus frog).

Eight young human adults (students from BIOL 315 Honor's Ecology section) were divided into three groups consisting of two or three members. Each group downloaded two unique sets of recordings used for this study. A process of validation was then conducted to develop a corrected dataset.

Trial II—ADRS were placed in a total of twenty locations along the banks of the Green River in south central Kentucky in 2007. The ADRS, the Amphibulator (Bowker and Cambron 2006), were placed on trees 3 m above the ground, recording for three minutes every hour. Recorders were clustered in three similar areas including Munfordville (Figure 2.3), Greensburg (Figure 2.4), and the UGRBP (Figure 2.5). Recorders represented similar habitats within 30 km of each other, but different microhabitats within the river corridor such as sloughs were near the recorders. A total of nine recorders, three from each area, were selected for the study due to the availability of uncorrupted recordings (Table 2.2). Recordings were selected and downloaded from May 24–28, 2007 due to the availability of recordings from all ADRS and the expected variety of anuran vocalizations in this period. Unlike the alternating hours of the first data set, all hours of the day were used for this data set. The eight species of interest chosen from this data set included: *B. americanus*, *B. fowleri*, *P. crucifer*, *P. feriarum*, *Lithobates (Rana) clamitans* (green frog), *Lithobates (Rana) catesbeiana* (bullfrog), *Hyla chrysoscelis* (Cope's gray treefrog), and *Acris crepitans* (cricket frog).

TABLE 2.2 A summary of the general area where groups of digital recorders were located (Munfordville, Greensburg, and the Upper Green River Biological Preserve), the number of the recorder, and coordinates of the recorder location in south central, KY.

Area	Recorder	Coordinates
Munfordville	1	N 37 18.643 W 85 46.125
Munfordville	3	N 37 17.699 W 85 46.802
Munfordville	5	N 37 17.429 W 85 47.621
Greensburg	2	N 37 16.038 W 85 34.902
Greensburg	8	N 37 16.139 W 85 35.709
Greensburg	10	N 37 16.162 W 85 36.200
UGRBP	11	N 37 14.880 W 85 59.193
UGRBP	15	N 37 14.610 W 86 00.274
UGRBP	19	N 37 14.668 W 86 01.218

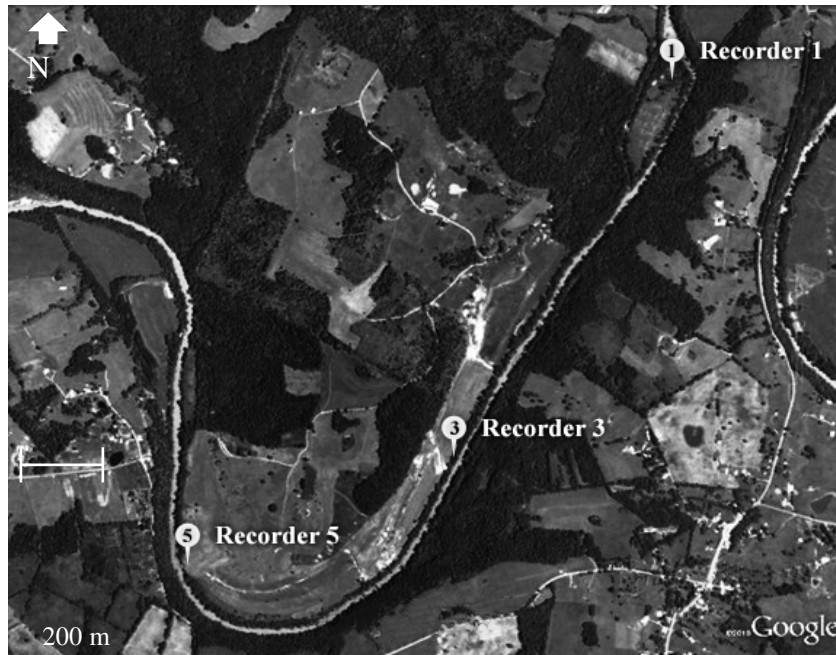


FIGURE 2.3 Satellite image of digital recorder locations number 1, 3, and 5 in the Munfordville area in south central Kentucky. Image retrieved from Google Maps.

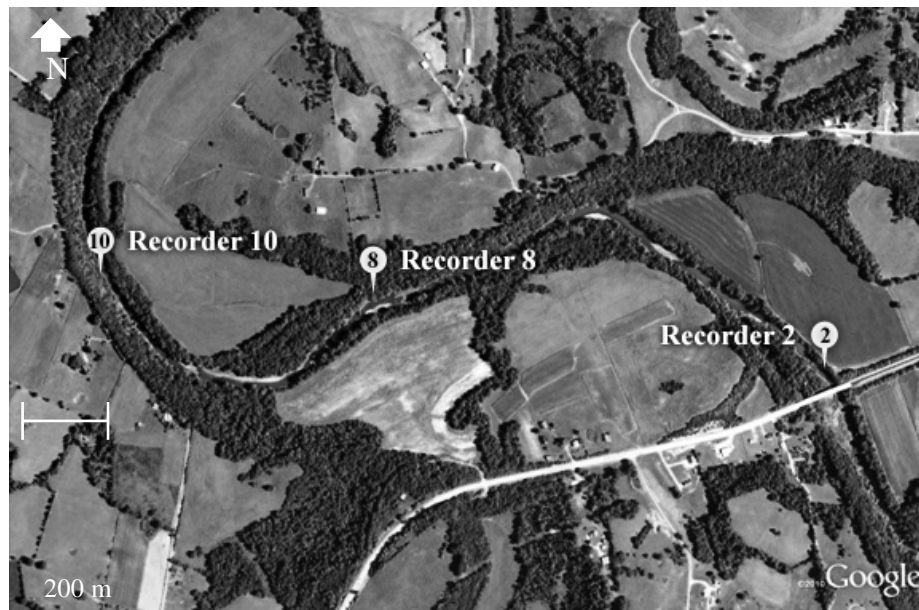


FIGURE 2.4 Satellite image of digital recorder locations number 2, 8, and 10 in the Munfordville area in south central Kentucky. Image retrieved from Google Maps.

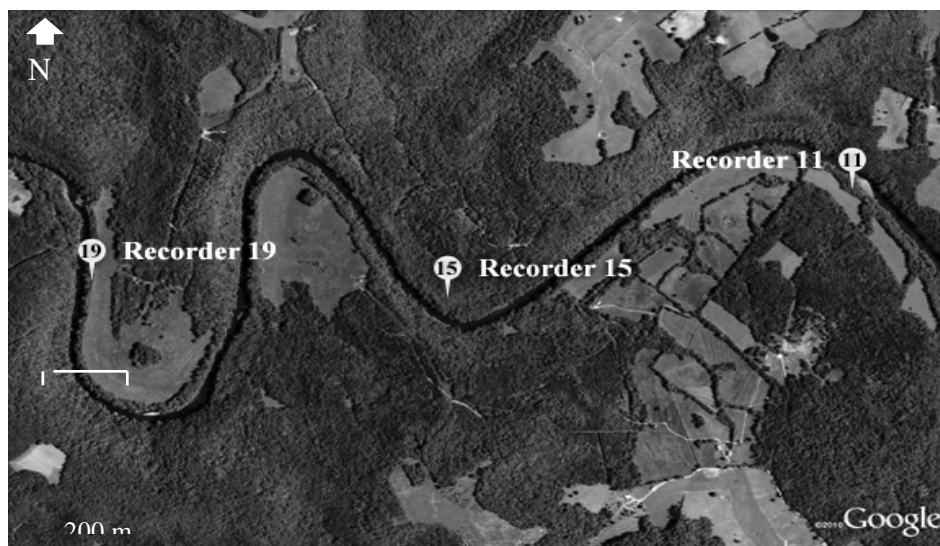


FIGURE 2.5 Satellite image of digital recorder locations number 11, 15, and 19 in the Upper Green River Biological Preserve in south central Kentucky. Image retrieved from Google Maps.

Eleven young human adults (students from BIOL 315 Honor's Ecology section) were divided into four groups of two or three listeners. Each group downloaded data from two recorders from different locations. All groups also downloaded one-fourth of the recordings from a single recorder (recorder 10).

Study Methods—In both trials, all listeners were trained during two sessions to identify the species of interest in each respective trial. Listeners were trained using both commercially available audio files and recordings from the ADRS. The listeners of the study had little previous experience in anuran ecology and no experience in the identification of their vocalizations. All listeners used the free audio program Audacity (v. 1.3.13 Beta). Each listener calibrated Audacity to identical settings of the spectrogram view that focused on the frequency range of anuran vocalizations. The listeners then viewed and listened to their assigned audio tracks on personal computers, primarily documenting the presence and absence of the target species for each 3-minute

recording. They also documented the number of calls when possible or otherwise marked “chorus” when too many individuals were calling to count. All listeners were provided with reference anuran vocalization recordings as well as recommended websites that they could use to help distinguish vocalizations. They could replay recordings and check vocalizations with reference vocalizations at their choosing. After group members finished their initial data collection, they collaborated within their assigned group to compare data, validate discrepancies, and compile a revised final data set.

I used the commercially available AVRS Song Scope[®] (Ver.3.1; Wildlife Acoustics Inc., Concord, Massachusetts, USA) to develop recognizer models to automatically detect anuran calls. A recognizer was developed for *A. crepitans*, *B. americanus*, *B. fowleri*, *H. chrysoscelis*, *L. catesbeiana*, *L. clamitans*, *P. crucifer*, and *P. feriarum*. The recognizers were then used to detect vocalizations for species in data sets that were also being evaluated by the human listeners. I validated results from the recognizers to remove false positive detections and develop a revised data set, noting the presence of errors. Results from the AVRS results and the human listeners were then compared. When results from the data sets did not match, I reviewed the track and noted the actual presence or absence of the vocalization. Using this method, I was able to produce a list of the actual presence and absence of each species during each track. Furthermore, I developed a list of the false positives and negatives produced by human listeners and the AVRS.

Statistical Methods—Because the data were based on frequency counts, a Wilcoxon signed-rank test was performed in SPSS (PSAW Statistics 2009) to detect differences between false positive and false negative detections for both human and

computer methods. The rate of false positive and false negative detections was also compared between human and computer methods.

I used contingency tables to perform a series of Chi-square tests in Microsoft Excel 2007 to determine the influence of each of the treatments (location, listener group, species, season, and recorder) on error rates of both human and computer methods. Because the number of false negatives was dependent upon the number of recordings with a species present, the error was considered in proportion to the total possible error for each comparison. For example, false positives were limited by the quantity of recordings where the species was not present. Therefore, the number of recordings without the species present was factored in proportion to false positive detections. False negative detections were limited by the actual presences of a species in a recording, thus the actual presence was considered in proportion to false negatives. Human error rates were considered to be independent of computer error rates, but false negatives and false positives in each computer and human datasets were not considered to be independent. A total of eight Chi-square tests were used for each dataset (two error types and four treatments), thus a Bonferroni correction was used to lower the acceptance of significant results to $\alpha=0.00625$. Human false negatives were not evaluated in Trial I due to the lack of this type of error ($n=1$), but human error probabilities were still penalized by the Bonferroni correction. Error rates were further evaluated by graphing the proportion similar to that considered in the contingency Tables with the following formulas:

Percent False Positives= number of false positive detections/ (number of recordings without the presence of a species)

Percent False Negatives= number of false negative detections/ number of actual presences of a species

RESULTS

Trial 1— A total of 793 (2,379 min.) recordings were evaluated by both human and computer detection methods. Of the 793 recordings, 55 were determined unusable due to recorder malfunctioning, leaving 737 recordings or 2,211 minutes of sound with complete evaluation performed.

Mean ranks of computer false positives were significantly different than false negatives ($W=434.5$, $P=0.002$). The true mean of computer false positives per species for each recorder ($n=5.42$) was higher than false negatives ($n=2.08$). Ranked human false positives were also significantly different than false negatives ($W=474$, $P<0.0001$). Actual means of human false positives were higher ($n=6.2$) than mean false negatives (0.04). There was no significant difference in ranked means of computer and human false positives ($W=230$, $P=0.23$). Ranked computer and human false negatives were significantly different ($W=188.5$, $P=0.004$).

After the validation process, human error was observed as 149 false positive detections and only one false negative detection. The majority of false positives were from *B. americanus* detections, and the sole false negative detection was of *P. crucifer* (Table 2.3). Computer error consisted of 130 false positive detections and 63 false negative detections. *Pseudacris feriarum* was responsible for the highest amount of false positive detections, and *P. crucifer* had the highest number of false negative detections (Table 2.3).

TABLE 2.3 Number of detections before validation for human listeners and computer (automatic vocalization recognition software) methods, number of actual presences, and number of recordings with detected error by error type for each species. HUM detect= number of human detections before validation, COM detect= number of computer detections before validation, HUMFP= human false positives, HUMFN= human false negatives, COMFP=computer false positives, COMFN=computer false negatives

Species	HUM detect	COM detect	Actual Presences	HUMFP	HUMFN	COMFP	COMFN
<i>B.americanus</i>	154	50	45	111	0	23	19
<i>B. fowleri</i>	12	30	10	2	0	24	4
<i>P. crucifer</i>	117	116	89	26	1	41	29
<i>P. feriarum</i>	30	51	20	10	0	42	11

Pseudacris crucifer was recorded more often than any other species and had the highest number of occurrences in April (Table 2.4). *Bufo americanus* was the next most prevalent species present with calling activity also largely during April, and *P. feriarum* was present only during April (Table 2.4). *Bufo fowleri* was the least recorded species and was only detected in June (Table 2.4). The only presence of a species in October was a single detection of *P. crucifer* (Table 2.4).

TABLE 2.4 Total number of 3min digital recordings with the presence of the four species evaluated in the study during the month that they were detected.

Species	April Presences	June Presences	October Presences
<i>B. americanus</i>	39	6	0
<i>B. fowleri</i>	0	10	0
<i>P. crucifer</i>	86	2	1
<i>P. feriarum</i>	20	0	0

Treatments Influencing Human Error in Trial I—Before validation, the human listeners recorded the total presence of the four species in Trial I similarly to the AVRS. However, a distinct difference was seen as humans falsely noted the presence of *B. americanus* more than the AVRS (Table 2.5).

Human false positives were influenced more by the evaluated treatments than any other error rate. Treatments causing disproportionate variation included human false positives by listener group, species, and season (Table 2.5). One group of listeners had a relatively lower false positive rate than other listener groups, 1.42%, compared to 6.09% and 7.90% (Figure 2.6). The percentages of possible human false positives were highest for *B. americanus* (15.99%) and followed by *P. crucifer* (4.00%), *P. feriarum* (1.39%), and *B. fowleri* (0.27%) (Figure 2.6). Human false positive rates by season were less variable, but June had less false positives (3.25%) than either April (5.60%) or October (5.33%) (Figure 2.7).

TABLE 2.5 Results of contingency table Chi-square tests that tested if the proportion of error rates separated among each Treatment (location, listener group, species or season) was different than expected. HUMFP= human false positives, COMFP=computer false positives, COMFN=computer false negatives, (*) denotes significance with Bonferroni correction of $\alpha=0.00625$

Treatment		HUMFP	COMFP	COMFN
Location	χ^2	0.539	9.219	0.447
	df	1	1	1
	<i>P</i>	0.463	0.0024*	0.504
Listener Group	χ^2	41.186	6.098	2.514
	df	2	2	2
	<i>P</i>	<0.0001*	0.0474	0.284
Species	χ^2	185.034	11.05	8.057
	df	3	3	3
	<i>P</i>	<0.0001*	0.0114	0.0449
Season	χ^2	20.506	6.523	6.366
	df	2	2	2
	<i>P</i>	<0.0001*	0.0383	0.0415

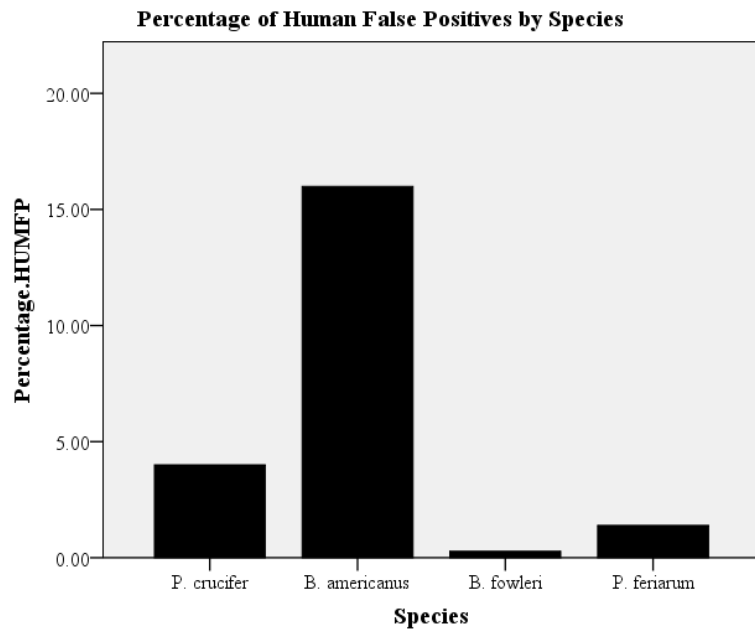


FIGURE 2.6 The percentage of false positive detections by humans according to each species. Percentage HUMFP= percentage human false positive detections

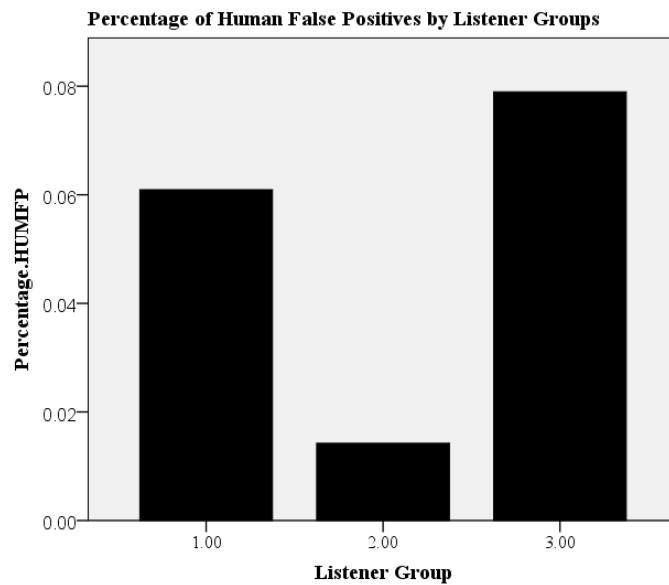


FIGURE 2.7 The percentage of false positive detections by human listener group 1, 2, and 3. Percentage HUMFP= percentage human false positive detections

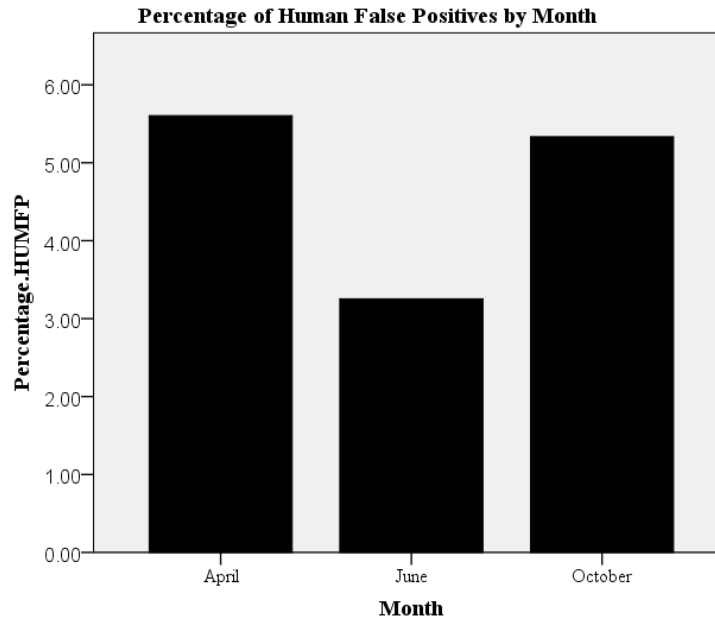


FIGURE 2.8 The percentage of false positive detections by humans for each month evaluated in Trial I. Percentage HUMFP= percentage human false positive detections

Treatments Influencing Computer Error in Trial I—Computer error occurred at a higher rate than human error but varied less according to the treatments of the study. Location was the only treatment that caused the AVRS false positives to vary significantly from the expected distribution (Table 2.7). Computer false positive detections occurred at a higher rate (5.98%) at the UGRBP than at WPP (3.44%) (Figure 2.9). The proportion of computer false negatives did not vary significantly from the expected distribution (Table 2.6).

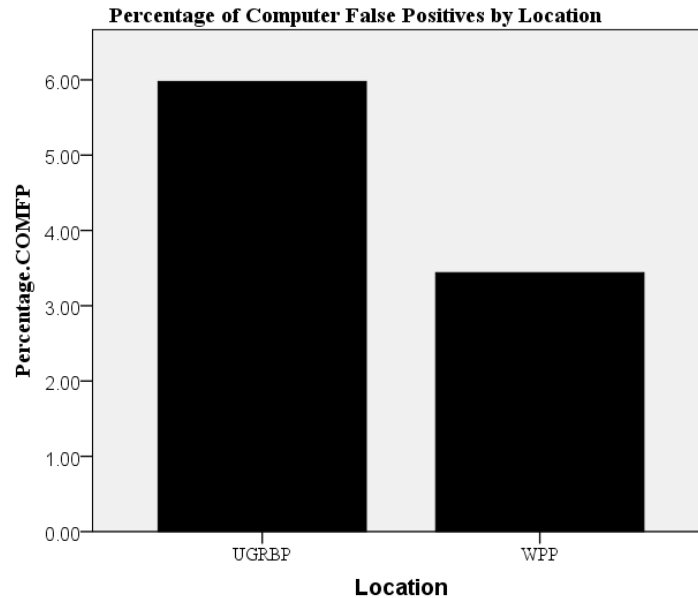


FIGURE 2.9 The percentage of computer false positive detections at the digital recorder locations at the Upper Green River Biological Preserve (UGRBP) and Weldon Peete Park (WPP) in Kentucky. Percentage COMFP= percentage of AVRS model false positives

Trial II—A total of 1080 recordings were evaluated by human and computer monitoring methods in Trial II. None of the recordings were unusable allowing the full 3,240 minutes of sound to be evaluated in this study.

A significant difference was found between the ranked mean of false positive and false negative error using computer methods ($W=3984.5$, $P<0.0001$). The true mean of false positives per species for each recorder ($n=14.91$) was higher than the mean of false negatives ($n=5.20$). Ranked human false positive and negative means were also significantly different ($W=3719$, $P<0.0001$). The true mean of false positives were also higher for humans ($n=3.64$) than false negatives ($n=0.072$). Computer false positive mean ranks were significantly different than human false positive ranks ($W=1030$, $p<0.0001$) as were false negative mean ranks ($W=1700.5$, $P<0.0001$).

Before validation of the results, human and computer detections were different in most cases (Table 2.6). Computer detections were higher for all species except for *B. americanus* and *B. fowleri* (Table 2.6).

After validation, *H. chrysoscelis* was found to have the most presences in recordings (Table 2.6). Very low presences were found for three of the four species found in Trial I, *B. americanus*, *P. crucifer*, and *P. feriarum* (Table 2.6).

Human error totaled 262 false positives and 52 false negatives for all species. *Bufo americanus* was the source of the most human false positives, and *H. chrysoscelis* caused the most human false negatives (Table 2.6). The error of the AVRS recognizer programs totaled 1074 false positives and 374 false negatives. The AVRS had the most false positive detections of *H. chrysoscelis* and the most false negative errors with *A. crepitans* (Table 2.6).

TABLE 2.6 Number of detections before validation for human and computer methods, number of actual presences, and number of recordings with error by error type for each species. HUM detect= number of human detections before validation, COM detect= number of computer detections before validation, HUMFP= human false positives, HUMFN= human false negatives, COMFP=computer false positives, COMFN=computer false negatives

Species	HUM detect	COM detect	Actual Presences	HUMFP	HUMFN	COMFP	COMFN
<i>A. crepitans</i>	132	248	132	12	12	212	97
<i>B. americanus</i>	96	12	4	93	1	11	3
<i>B. fowleri</i>	174	114	145	29	1	29	60
<i>H. chrysoscelis</i>	198	480	174	55	30	394	90
<i>L. catesbeiana</i>	178	213	136	42	0	151	75
<i>L. clamitans</i>	78	190	58	28	8	180	48
<i>P. crucifer</i>	3	76	1	2	0	76	1
<i>P. feriarum</i>	5	21	0	5	0	21	0

Treatments Influencing Human Error in Trial II—Overall, errors in Trial II were more influenced by the evaluated treatments than those in Trial I. Human error was again more influenced by the studied treatments than computer error (Table 2.7). Human false positive rates were significantly different from expected by all treatments evaluated in Trial II (location, listener group, species, and recorder) (Table 2.7). Human false positive rates were moderately higher at Munfordville (4.72%), followed by the UGRBP (3.17%) and Greensburg (1.85%) (Figure 2.10). The rate of human false positives was lowest for listener group 2 (1.48%) and group 3 (0.81%), and these rates were highest for listener group 1 (6.68%) and group 4 (3.99%) (Figure 2.11). Human false positive rates were highest for *B. americanus* (8.64%) and *H. chrysoscelis* (6.07%), moderate for *L. catesbeiana* (4.45%), *B. fowleri* (3.10%), and *L. clamitans* (2.24%), and lowest for *A. crepitans* (0.84%), *P. feriarum* (0.46%), and *P. crucifer* (0.19%) (Figure 2.12). Human false positive rates also varied significantly from the expected distribution depending upon the recorder, ranging from 0.25–8.00% (Figure 2.13)

TABLE 2.7 Results of Chi-square tests that tested if the proportion of error rates separated among each treatment (location, listener group, species or season) was different than expected. HUMFP= human false positives, COMFP=computer false positives, COMFN=computer false negatives, (*) denotes significance with Bonferroni correction of $\alpha=0.00625$

Treatment		HUMFP	HUMFN	COMFP	COMFN
Location	χ^2	32.01	0.705	1.82	12.37
	df	2	2	2	2
	<i>P</i>	<0.0001*	0.703	0.401	0.002*
Listener Group	χ^2	125.28	13.61	11.94	3.36
	df	3	3	3	3
	p-value	<0.0001*	0.0035*	0.0076	0.340
Species	χ^2	186.38	39.04	814.37	12.29
	df	7	4	7	4
	<i>P</i>	<0.0001*	<0.0001*	<0.0001*	0.0153
Recorder	χ^2	216.34	31.31	143.36	25.01
	df	8	8	8	8
	<i>P</i>	<0.0001*	<0.0001*	<0.0001*	0.0015*

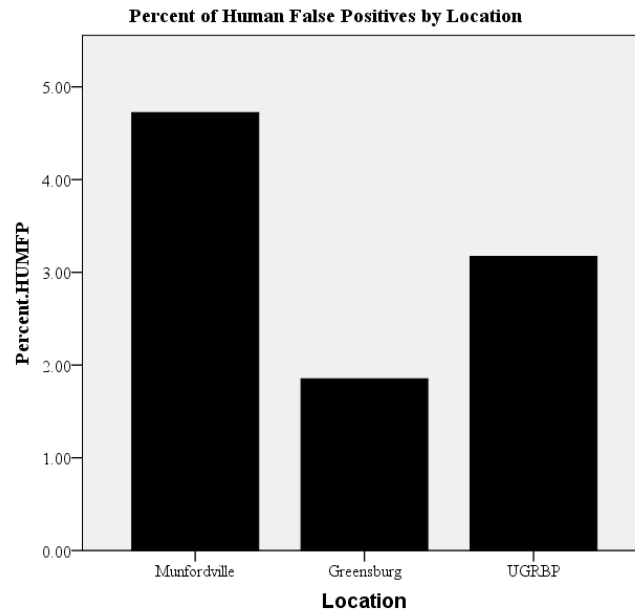


FIGURE 2.10 The percentage false positive detections by human listeners in each area containing recorders (n=3). Percentage HUMFP= percentage human false positives

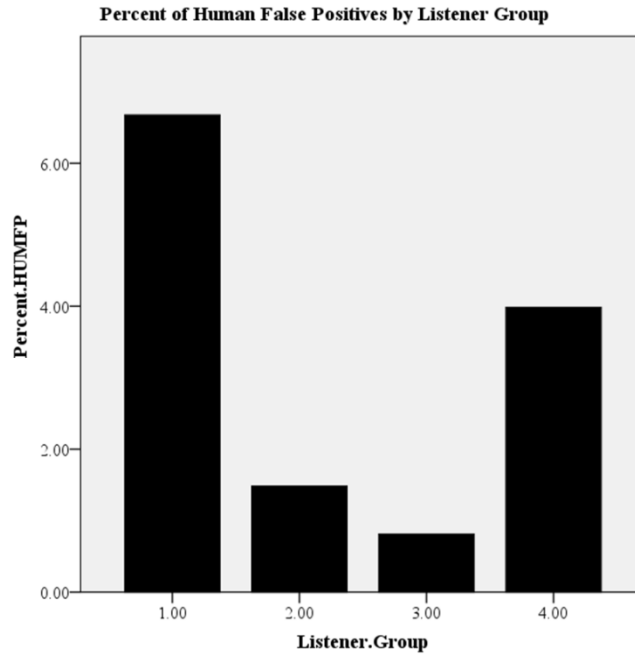


FIGURE 2.11 The percentage of false positive detections by the human listener groups 1, 2, 3, and 4. Percentage HUMFP= percentage human false positive detections

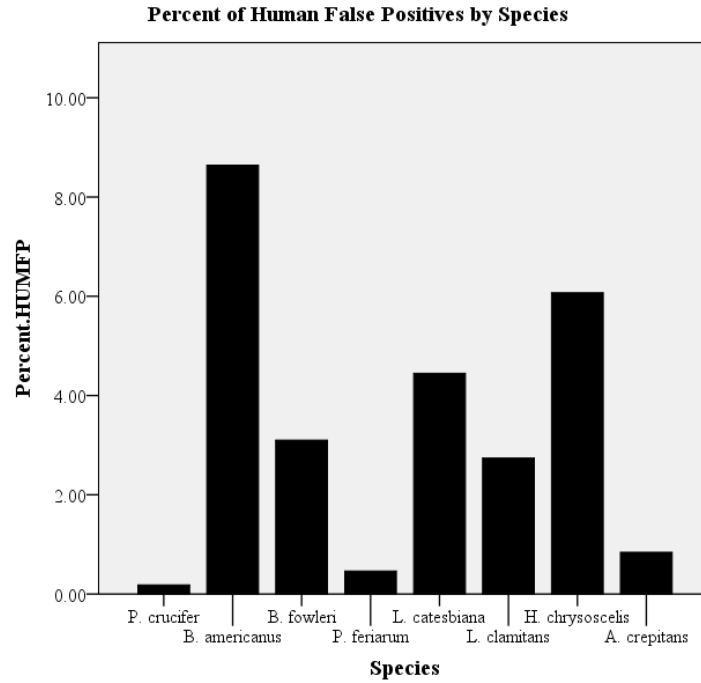


FIGURE 2.12 The percentage of false positive detections by humans for each species evaluated. Percentage HUMFP= percentage human false positive detections

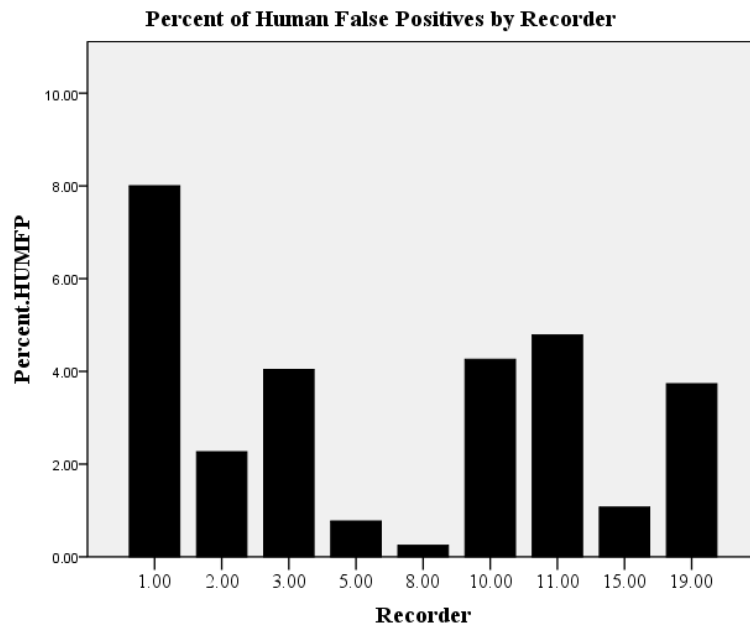


FIGURE 2.13 The percentage of human false positive detections by each recorder used in Trial II. Percentage HUMFP= percentage human false positive detections

Species with low possible false negatives due to few actual presences, including *B. americanus* (n=4), *P. crucifer* (n=1), and *P. feriarum* (n=0), were excluded from analyses of false negatives with species as a treatment for computer and human error. Human false negatives varied from the expected proportions according to species and recorder but not location or listener group. Of the five species evaluated, the highest human false negative rates resulted from *H. chrysoscelis* (17.24%), *L. clamitans* (13.79%), and *A. crepitans* (9.09%) whereas the lowest rates were from *B. fowleri* (0.69%) and *L. catesbeiana* (0.00%) (Figure 2.14). Listener groups had a high degree of variance from the expected distribution in their percentage of false negative detections, ranging from 0.81% to 12.89%. Human false negative also varied from the expected distribution when analyzed according to recorder, ranging from 0.00–20.59% (Figure 2.15).

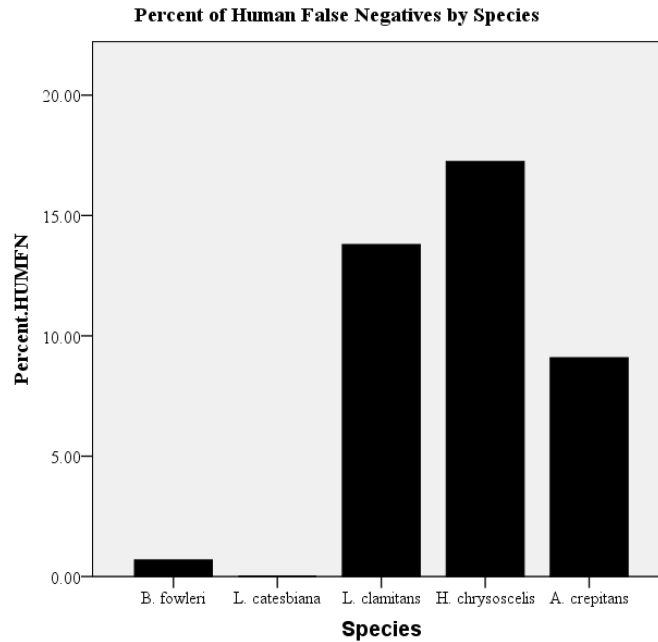


FIGURE 2.14 The percentage of false negative detections by humans for each species with greater than 5 detections. Percentage HUMFN= percentage human false negative detections

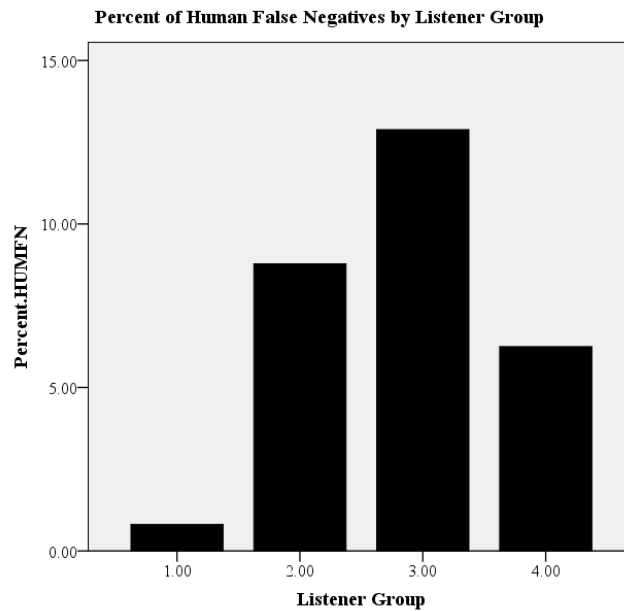


FIGURE 2.15 The percentage of human false negative detections by listener groups 1, 2, 3, and 4. Percentage HUMFN= percentage human false negative detections

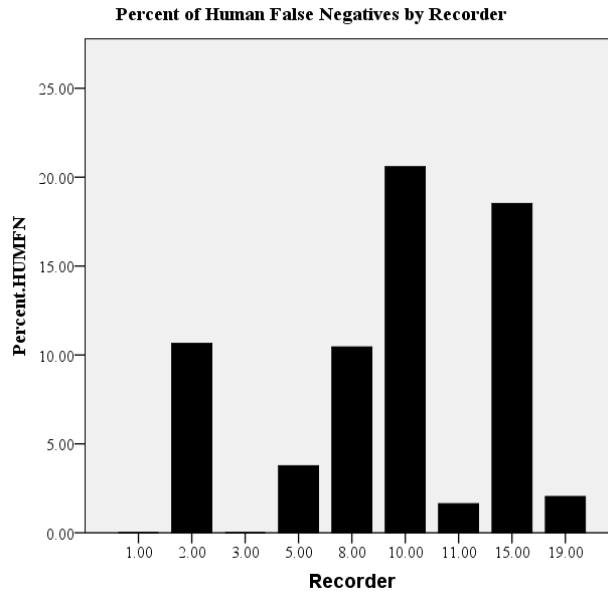


FIGURE 2.16 The percentage of human false positive detections by each recorder used in Trial II. Percentage HUMFN= percentage human false negative detections

Treatments Influencing Computer Error in Trial II—The rate of false positives generated by the AVRS in Trial II significantly differed from the expected proportion by species and recorder (Table 2.8). Species with the highest rate of false positive detection included *H. chrysoscelis* (43.49%), *A. crepitans* (22.36%), *L. clamitans* (17.61%), and *L. catesbeiana* (16.00%) whereas species with lower rates of false positives included *P. crucifer* (7.04%), *B. fowleri* (3.10%), *P. feriarum* (1.94%), and *B. americanus* (1.02%) (Figure 2.17). False positive rates evaluated by each recorder ranged from 10.63–16.58% (Figure 2.18).

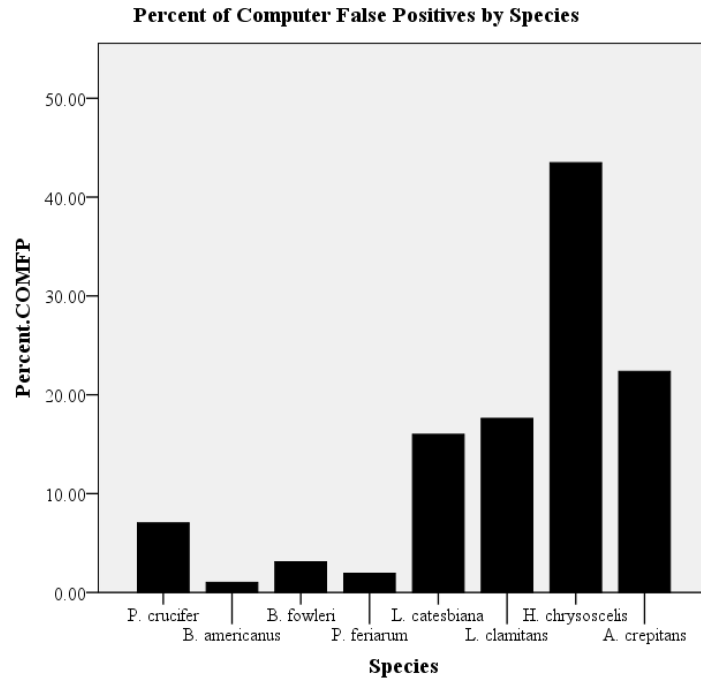


FIGURE 2.17 The percentage of computer false positive detections for each species evaluated. Percentage COMFP= percentage computer false positive detections

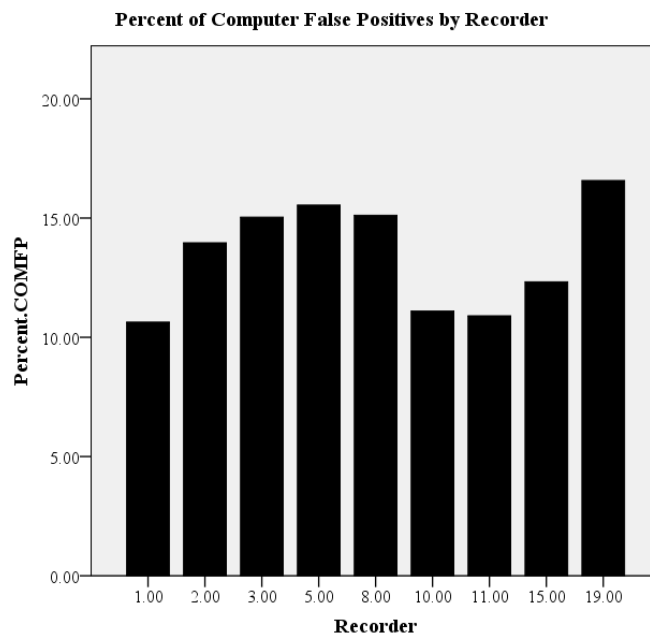


FIGURE 2.18 The percentage of AVRS false positive detections by each digital recorder (n=9). Percentage COMFP= percentage computer false positive detections

Of the treatments in Trial II, location and recorder (a subset of location) caused computer false negative rates to significantly vary from the expected (Table 2.8). The highest false negative rates occurred at Munfordville (79.41%) followed by Greensburg (52.48%) and the UGRBP (43.07%) (Figure 2.19). False negative rates by individual recorders were much higher than that of human listeners and varied from 32.79% (recorder 19 at the UGRBP) to 91.67% (recorder 1 at Munfordville) (Figure 2.20).

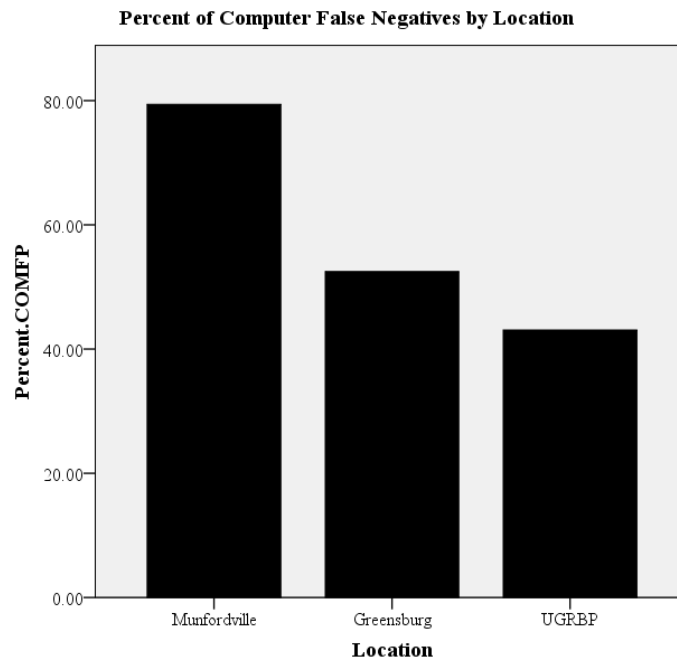


FIGURE 2.19 The percentage of computer false negative detections by the digital recorders (n=3) in each area (Munfordville, Greensburg, and the Upper Green River Biological Preserve (UGRBP)). Percentage COMFN= percentage computer false negative detections

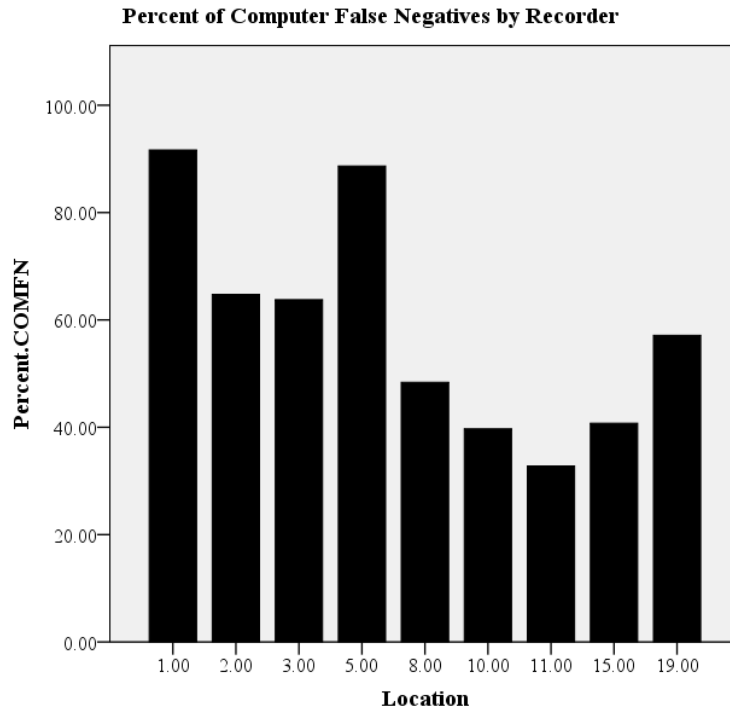


FIGURE 2.20 The percentage of computer false negative detections each of the digital recorders. Percentage COMFN= percentage computer false negative detections

DISCUSSION

This study highlighted several key influences on error rates that should be considered in an effective anuran monitoring program. As hypothesized, human error was homogeneous and random. However, unlike my hypothesis, computer error was also not homogeneous and random. By and large, computer error occurred at higher rates, but human error was more likely to be influenced by the treatments (species, location, listener group, season, and recorder) evaluated in the study. The detection of false positives experienced more biased variation whereas false negatives were less likely to be influenced by the treatments. Among the treatments, species was the most influential treatment in causing varied error rates. In addition to species, the listener groups were a source of variance in the human error rates but not with the AVRS as would be expected.

Finally, the location and the individual recorders were the greatest influencing treatments on AVRS error rates, and species was a strong influence in Trial II. The results supported the hypotheses that human error is homogeneous and random and did not support the hypothesis that computer error is homogeneous and random.

Computer Error—The development of a recognizer model in Song Scope is a time-consuming and moderately subjective task. The creation and modification of a recognizer for a given species required approximately 15 h to 40 h for each species. The time I spent on construction of recognizer models is similar to that of Waddle et al. (2009). The problem faced in adjusting a recognizer model is twofold. First, a recognizer program must account for variation within a species' calls, yet still be specific enough to reduce extraneous false positives. This task can be especially complicated given the variation in the frequency, structure, and length of anuran vocalizations. Second, a recognizer program must have an acceptable balance of error rates. To achieve this goal, one must decide the desired balance of error rates by adjusting the score parameters available in Song Scope. A lower score setting will likely reduce false negatives but also increase false positives. The opposite effect occurs when scores are adjusted higher. A prior study faced the same problem with Song Scope, and models in their study resulted in higher false negative than false positives (Waddle et al. 2009). In my study, I attempted to reduce false negative errors without unreasonably increasing false positives. The rationale behind this decision was that false positive detections can be virtually eliminated relatively quickly by manually validating the results generated by the AVRS. This decision was supported because computer error did produce higher amounts of false positives than false negatives.

It is not surprising that error rates were higher with the AVRS than human listeners. Because the AVRS is based on digital information, a variety of factors can cause errors that are not problematic for human discernment. Humans have also been observed to be more accurate than AVRS in the detection of bats (Skowronski and Fenton 2008). For example, the “click” vocalization of *Acris crepitans* is usually repeated in a series that is quite variable in duration and repetitions. Considering that, calls are made singly at times, and the full series of repetitions is often disrupted by other sounds, a recognizer for this species consisted of only one “click”. This vocalization is similar to background noises in the same environment such as branches snapping or rain pattering on tree leaves that lead to an increased false positive rate (22.36%). Automatic detection of bird vocalizations including only a single syllable has also been shown to increase error rates (Somervuo et al. 2006). The *A. crepitans* call at a distance was also often low enough in amplitude to not be a sound “considered” by the AVRS leading to a high rate of false negatives (73.48%). The problem of distant and quiet calls was also experienced with other species. In contrast, species with higher amplitude calls in frequency ranges that were less likely to be competing with other sounds such as other louder chorusing species (e.g. *P. crucifer*) or insects were better detected by AVRS.

The probable impact of call amplitude on error rates emphasizes the need to place ADRS as closely as possible to breeding locations where calls can clearly be heard. This study was unique in that ADRS were placed along river corridors and recorded during all hours of the day instead of the usual placement in heavily used breeding locations such as wetlands or ponds, often with recordings only during night. This arrangement likely

increased error rates as fewer vocalizations were present than could potentially be found in more ideal conditions for anuran reproduction.

Computer false positives were affected only by location (synonymous with recorder) in Trial I and by species and recorder in Trial II. The false positive rates in Trial II were almost twice as high at the UGRBP than WPP. This was unexpected given the greater diversity of noise present in a more urban habitat, suggesting that the type of noise present is less influential than specific sounds. Background noise has previously been shown to have an effect on AVRS in the error rates in detection of birds (Trifa 2008) and bats (Skowronski and Fenton 2008). For example, the greater amount of insect noise at the UGRBP that may explain higher false positives rates in Trial I. In Trial II, the species with the lowest false positive error rates were *Bufo americanus*, *Bufo fowleri*, and *Pseudacris feriarum*. *Pseudacris crucifer* had a moderate amount of false positive detections. Among these species, *B. fowleri* was found in much higher quantities (n=145) yet produced a proportionately low a number of false positives (3.10%). These species potentially had the most distinct vocalizations from background noises with fewer call variations, with the exception of *B. americanus*.

Species with computer false positive detection rates much higher than other species included *Acris crepitans*, *Hyla chrysoscelis*, *Lithobates catesbeiana*, and *Lithobates clamitans*. Although it is uncertain why false positive rates of these vocalizations were higher than in other species, the development of an accurate recognizer model for these species proved to be more difficult. The most prevalent species in the recordings was *H. chrysoscelis*. Birds calling in a similar frequency range of *H. chrysoscelis* appeared to be the primary source of noted error. *Lithobates*

catesbeiana detections were highly affected by background noise found in low frequencies such as wind, cattle, and vehicles (both airplanes and automobiles). Calls of *L. clamitans* were misidentified by the AVRS at high rates compared to actual presences, and false positives often resulted from background noises such as those affecting *A. crepitans*. Additionally, river water noise and low frequency bird calls were commonly subject to AVRS false detections. Interestingly, *L. clamitans* also had the highest rate of errors in the evaluation of Song Scope by Waddle et al. (2009) in a different season and location. This suggests that some species are likely less suited for accurate identification with AVRS, a trend also seen with birds (Somervuo 2006). Fortunately, false positives of the AVRS can be validated quickly for moderate numbers of false positive results, and a skilled listener should be able to virtually eliminate false positive errors from AVRS detections.

Of greater concern with the AVRS is the rate of false negative detections. Recorder and location was a source of false negative variation in Trial II. It is not surprising that these treatments interrelate as three recorders were placed in each location as part of the study methods. The recorder with the highest error was at recorder 1 at Munfordville (91.76%) at a site noted for distant calling activity. This supports the hypothesis that distant, low amplitude calls are less likely to be detected, a problem previously noted for Song Scope (Agranat, I. 2007. *Automatic detection of cerulean warblers using autonomous recording units and Song Scope Bioacoustics Software*. Available from wildlifeacoustics.com [accessed 28 November 2011]). Therefore, false negatives were again probably due to distance from the breeding site, emphasizing the need for a close proximity to a breeding location. Unlike false positives in AVRS, false

negative rates can only be reduced by reconfiguration of the recognizer model or, to an extent, by lowering score settings, leading to higher rates of false positives.

Consequently, perhaps the best solution to reduce false negatives is to lower the recognizer model settings (within reason) and change recorder location if possible.

Human Error—The sources of human error are more difficult to evaluate. This is primarily because in a study with multiple listeners, it is difficult to discern the cause of error is at the time of detection. I validated results by listening to recordings in question, only enabling me to note whether species were present or absent as the listener had suggested, although probable causes were noted. Overall, humans showed the potential to reduce error rates but were more affected by the studied treatments.

As with the construction of a recognizer model, time is required to train human listeners to better identify calls. The listeners in this study were trained as a group for no longer than three hours and were instructed to learn the anuran calls required for the study on their own time. The most time-consuming portion of the study for listeners was the approximately 12 hours spent listening to recordings, plus the time listener groups spent validating recordings. Human listeners in this study were disadvantaged in that they were inexperienced in anuran call identification and did not know that their error rates were being evaluated. Conversely, they were highly advantaged in their ability to listen to example calls when questionable sounds were heard, to observe visual cues in spectrograms while listening, and to crosscheck with at least one other person to aid in validation of results. Given the use of multiple observers, the amount of observed human error was higher than expected.

Similar to computer error, humans were biased to have a greater number of false positives than false negatives. In general, error rates were also higher in Trial II. This suggests that identification of a greater number of species may cause higher error rates. Other explanations may account for these error rates, however, including the higher identification difficulty level of calls unique to Trial II, distance of calling individuals, and the overall greater number of calls. Unfortunately, false positive errors would likely go unnoticed in a traditional monitoring program with the exception of species inclusions during times outside of their normal breeding season. Therefore, the listener's observations would likely be accepted as "fact" when less precision should be awarded to their results.

The greatest source of variation in listener errors was species identification. In the Trial I, misidentification of *B. americanus* was the greatest source of error. Misidentification of this species was usually in recordings with insect vocalizations, many sounding similar to the long trill of *B. americanus*. False positives were particularly high when *Oecanthus* sp. (tree cricket) vocalizations were present in October recordings having a similar frequency and spectrogram appearance to *B. americanus* (Figure 2.21). A basic knowledge of the ecology of *B. americanus* would have allowed the listeners to know that this species breeds primarily in the early spring (Conant and Collins 1998). Therefore, as previously suggested, a greater emphasis should be placed on the basic biology and ecology of the species being identified (Genet and Sargent 2003). Consequently, identifications such as those of *B. americanus* in the fall should be treated with greater scrutiny by listeners. This increase in scrutiny would lead to decreased error.

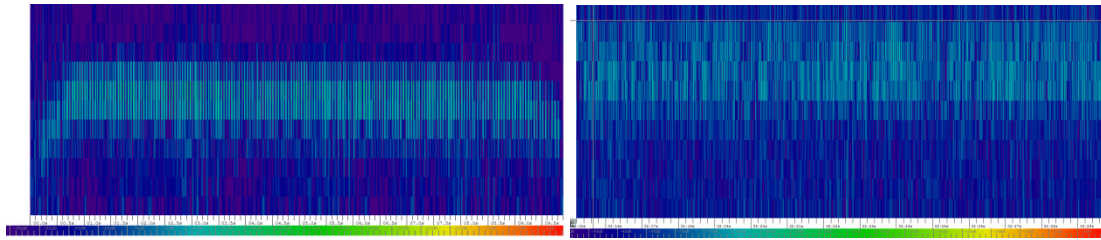


FIGURE 2.21 Spectrograms of *B. americanus* (left) and *Oecanthus* sp. (right) that were often confused by human listeners during October.

In trial II, even in species with limited or no detections, *P. crucifer* (n=1) and *P. feriarum* (n=0), were still subject to false positive identifications by humans. Detections of species with no actual presence (phantom species) are also noted by McClintock et al. (2010), presenting a danger of species inclusion of a species with no actual presence in a monitoring program. False positives for *A. crepitans* were also low with similar sounds of birds and insects appearing most often in misidentified tracks. Once again, false positives were highest for *B. americanus* even though this species had only four actual presences in the recordings. Most of the false positives for *B. americanus* were during recordings wherein insect sounds were prevalent. Moderately high amounts of false positive detections also occurred for *B. fowleri*, *H. chrysoscelis*, *L. catesbeiana*, and *L. clamitans*. The misidentification of *B. fowleri* was surprising given results from Trial I and the distinctiveness of its vocalization. Most causes for misidentification of this species were unknown, although a few dog, cow, and other anuran vocalizations were noted in misjudged recordings. Because actual presences were much higher in Trial II than Trial I, it is likely that listeners became accustomed to actively searching for this vocalization. The reason for the majority of *H. chrysoscelis* misidentifications could not be determined from validated recordings although some appeared to be caused by bird

vocalizations. Reasons for misidentification of *L. catesbeiana* and *L. clamitans* were generally not discernible from recordings.

Human false negatives also varied by species in Trial II. The evaluation of missed calls was difficult as it is difficult to judge why another individual did not notice the presence of a species. The highest proportions of false negatives were found for *H. chrysoscelis*, *A. crepitans*, and *L. clamitans*. Calls of these species are generally less distinct than those with lower false negative rates, *B. fowleri* and *B. catesbeiana*. Because the AVRS had an increased rate of false negatives and was used to detect species not recorded by human listeners, it is possible that additional false negative detections of human listeners were not identified by the study.

This study's results reflected the findings of a previous study wherein human listeners were evaluated for anuran identification in recordings (Lotz and Allen 2007). In both studies, human listeners misidentified species at different rates (Lotz and Allen 2007). However, unlike this study, their respondents were more likely to have false negative errors than false positives. This may be due to the other study methods where a small number of tracks (n=12) with several species present on each track was used, thus allowing for fewer errors (Lotz and Allen 2007). Variation in detection by species was also seen in other studies (Lotz and Allen 2007; de Solla et al. 2005). Overall, higher false positive than false negative rates in these studies were consistent with my results. As observed in another study, *Bufo americanus* was the source of high rates of false positives (Lotz and Allen 2007). *B. americanus* has also been subject to identification in tracks where it did not exist (phantom species) (McClintock et al. 2010). Unlike this

study, *Acris crepitans* was a species that had higher misidentification rates in other studies (Lotz and Allen 2007; McClintock, et al. 2010).

In a study, even small variation in error rates of different human observers could lead to considerable error in the estimation of population size and composition (Gooch et al. 2006). Additionally, even the fundamental presence or absence of a species is subject to misjudgment. Conservation managers should take detection probabilities and human error rates into consideration when determining whether a manual call survey program is appropriate for their goals of their monitoring program.

Variation in false positive detections by listeners according to their assigned listener group and recorder was a significant treatment in Trial II. It has been suggested that multiple observers in manual calling surveys would decrease error rates (Pierce and Gutzwiller 2007). Even with the implementation of groups in this study, variation in listener identification rates still existed. Unlike many other studies, in both Trial I and Trial II, the listeners held equal amounts of experience with call identification although they had a basic ecological background. Prior experience of an observer has largely not been observed to be a factor in identification error (Genet and Sargent 2003; Lotz and Allen 2007; Shirose et al. 1997), but a higher level of experience was present in all other cases. The conclusions of this study reflect the results of studies in that listener bias was both found to a significant factor (e.g. McClintock et al. 2010). It is also important to note that listener bias may be an artifact of sampling design. A listener group may have been assigned two recorders wherein conditions for identification were more difficult. False negatives and positive variation existed by recorder in Trial II, and it is probable that the same treatments mentioned for computer error also influenced listener error.

Prior studies have concluded that the distance of the call, background noises, and observer experience are important predictors of human error rates, all probable factors in this study (McClintock et al. 2010; Dorcas 2009)

Recommendations—It is to be expected that error will always be a problem even with the most skilled human observer and future advances in computer technology. Methods such as AVRS show potential to at least reduce bias that will be a persistent problem in manual call surveys. This study supports the suggestions to include multiple observers, screen potential participants, and assess participants for error rates to improve the effectiveness of manual calling surveys (McClintock et al. 2010). Moreover, as detection rates become known for species, they can be incorporated in statistical methods used to monitor population trends (de Solla et al. 2005). The low false negative rates of human listeners and the possibility of virtually eliminating false positives with AVRS makes the combination of the two methods for studies requiring highly accurate data appealing. Additionally, the cross-examination of these methods allows for a relative assessment of error rates that can be used in future studies. The cost and time associated with performing both methods simultaneously does make this solution less attractive.

The most appropriate method to use for anuran monitoring should be dependent upon the pre-determined goals of the study. Human listeners would likely be both time and cost effective for short-term studies that do not require the evaluation of copious amounts of auditory data. Specifically, manual call surveys would likely allow for the collection of small amounts of information from a wide variety of geographic locations. The primary drawback to human listeners as detected in this study is high false positive rates that would likely lead to biased and varied population estimates. Also, if traditional

manual call surveys were used without ADRS, a lower quality information would be provided than the hourly-scale, reviewable, and permanent auditory data of ADRS. This also decreases the ability to identify species with low detection probability.

The use of ADRS and AVRS would be best suited for long-term monitoring of specific locations or for collection of data for scientific purposes. The use of ADRS in a variety of locations is restricted by the current cost of ADRS. As technology progresses, cost will potentially become a less significant factor. Although startup cost and time involved with using ADRS and AVRS is fairly high, their use for a long-term study should be both time and cost effective. The development of recognizer models for all species is a concern and may initially require dependence on calls from other sources until enough quality vocalizations are found in the area to construct the needed recognizer models. Still, the ability to clearly visualize calls with a spectrogram and validate computer false positive detections is an advantage over human listeners. Unfortunately, the high number of false negatives causes the computer to lack the precision of human listeners. This shortcoming should be overcome by increasing the probability of detection through the employment of hourly-scale recording analysis.

The results of this study clearly show bias for both humans and computer methods in auditory monitoring, and these factors should be considered if anuran population trends are to be effectively monitored and therefore best conserved. Future studies should develop methods that allow for an accurate account of sources of error. It would also be of interest to incorporate a calling frequency estimate such as that used in the NAAMP (Weir and Mossman 2005).

BIBLIOGRAPHY

- Alford, R.A. and S.J. Richards. 1999. Global amphibian declines: a problem in applied ecology. *Annual review of ecology and systematics* 30:133–165.
- Bee, M.A., and E.M. Swanson. 2007. Auditory masking of anuran advertisement calls by road traffic noise. *Animal Behavior* 74:1765–1766.
- Blaustein, A.R., D.B. Wake, W.P. Sousa. 1994. Amphibian declines: judging stability, persistence, and susceptibility of populations to local and global extinctions. *Conservation Biology* 8:60–71.
- Brandes, T.S., P. Naskrecki, and H.K. Figueroa. 2006. Using image processing to detect and classify narrow-band cricket and frog calls. *Journal of the Acoustical Society of America* 120:2950–2957.
- Bridges, A.S. and M.E. Dorcas. 2000. Temporal variation in anuran calling behavior: implications for surveys and monitoring programs. *Copia* 2:587–592.
- Collins, J.P. and A. Storfer. 2003. Global Amphibian Declines: Sorting the Hypothesis. *Diversity and Distributions* 9:89–98.
- Conant, R., & Collins, J.T. 1998. *A Field Guide to the Reptiles and Amphibians of Eastern and Central North America*. Third Edition, Expanded. Houghton Mifflin Company. New York, New York, USA.
- de Solla, S.R., L.J. Shirose, K.J. Fernie, G.C. Barrett, C.S. Brousseau, and C.A. Bishop. 2005. Effect of sampling effort and species detectability on volunteer based anuran monitoring programs. *Biological Conservation* 121:585–594.
- Dorcas, M.E., Price, S.J., Walls, S.C., and Barichivich, W.J. 2009. Auditory Monitoring of Amphibian Populations. Pp. 281–298 *In* *Amphibian Ecology and Conservation: A Handbook of Techniques*. Dodd Jr., K. Oxford University Press, New York, New York, USA.
- Gerhardt, H.C. and F. Huber. 2002. Acoustic communication in insects and anurans: common problems and diverse solutions. University of Chicago Press, Chicago, Illinois, USA.
- Gerhardt, H.C. and M.A. Bee. 2006. Recognition and Localization of Acoustic Signals. Pp. 113–146 *In*, *Hearing and sound communication in amphibians*. P.M. Narins, A.S. Feng, R.R. Fay, and A.N. Popper. Springer-Verlag, New York, New York, USA.

- Gerhardt, H.C. and G.M. Klump. 1988. Phonotactic responses and selectivity of barking treefrogs (*Hyla gratiosa*) to chorus sounds. *Journal of Comparative Physiology* 163:795–802.
- Gibbs, J.P., K.K. Whiteleather, and F.W. Schueler. 2005. Changes in frog and toad populations over 30 years in New York state. *Ecological Applications* 15:1148–1157.
- Gooch, M.M., A.M. Heupel, S.J. Price, and M.E. Dorcas. 2006. The effects of survey protocol on detection probabilities and site occupancy estimates of summer breeding anurans. *Applied Herpetology* 3:129–142.
- Lotz, A. and C.R. Allen. 2007. Observer bias in anuran call surveys. *Journal of Wildlife Management* 71:675–679.
- Mazerolle, M.J. and L.L. Bailey, W.L. Kendall, J.A. Royle, S.J. Converse, and J.D. Nichols. 2007. Making great leaps forward: accounting for detectability in herpetological field studies. *Journal of Herpetology* 41:672–689.
- McClintock, B. T., L.L Bailey, K.H. Polluck, and T.R. Simons. 2010. Experimental investigation of observation error in anuran call surveys. *Journal of Wildlife Management* 74:1882–1893.
- PASW Statistics. 2009, July 30. SPSS statistical software package Release 18.0.0.
- Pierce, B.A., and K.J. Gutzwiller. 2007. Interobserver variation in frog call surveys. *Journal of Herpetology* 41:424–429.
- Royle, J.A. and W.A. Link. 2006. Generalized site occupancy models allowing for false positive and false negative errors. *Ecology* 87:835–841.
- Shirose, L.J., C.A. Bishop, D.M. Green, C.J. MacDonald, R.J. Brooks, and N.J. Helferty. 1997. Validation test of an amphibian call count survey technique in Ontario, Canada. *Herpetologica* 53:312–320.
- Skowronski, M.D. and M.B. Fenton. 2008. Quantifying bat call detection performance of humans and machines. *Journal of the Acoustical Society of America* 125:513–521.
- Somervuo, P., A. Harma, and S. Fagerlund. 2006. Parametric representation of bird sounds for automatic species recognition. *IEEE Transactions on Audio, Speech, and Language Processing* 14:2252–2263.

- Stuart, S.N., J.S. Chanson, N.A. Cox, B.E. Young, A.S. Rodrigues, D.L. Fischman, et al. 2004. Status and Trends of Amphibian Declines and Extinctions Worldwide. *Science* 306:1783–1786.
- Trifa, V.M., A.N. Kirschel, C.E. Taylor, and E.E. Vallejo. 2008. Automated species recognition of antbirds in a rainforest using hidden Markov models. *Journal of the Acoustical Society of America* 123:2424–2431.
- Trimboli, S. R. 2010. A comparison of the singing activity of the Carolina Wren (*Thryothorus Ludovicianus*) in Urban and Rural Settings. M.Sc. Thesis, Western Kentucky University, Bowling Green, KY, USA. 98 p.
- Waddle, J.H., T.F. Thigpin, and B.M. Glorioso. 2009. Efficacy of Automatic Vocalization Recognition. *Herpetological Conservation and Biology* 4:384–388.
- Weir, L.A. and M.J. Mossman. 2005. North American Amphibian Monitoring Program (NAAMP). Pp. 307–313 *In* Amphibian Declines: Conservation Status of United States Species, Lannoo, M. J. University of California Press, Berkeley, California, USA.
- Weir, L.A., J.A. Royle, P. Nanjappa, and R.E. Jung. 2005. Modeling anuran detection and site occupancy on the North American Amphibian Monitoring Program (NAAMP) routes in Maryland. *Journal of Herpetology* 39:627–639.
- Wells, K.D. and J.J. Schwartz. 2006. The behavioral ecology of anuran communication. *In* Hearing and sound communication in amphibians, Narins, P. M., A. S. Feng, R. R. Fay, & A. N. Popper (pp. 44–86). Springer-Verlag, New York, New York, USA.

